GenCore version 5.1.6


```
OM nucleic - nucleic search, using sw model

Run on:         June 10, 2003, 03:35:56 ; Search time 5459 Seconds
                                          (without alignments)
                                          11579.281 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  2172
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2172	100.0	2224	9	AY004154	AY004154 Homo sapi
2	2164	99.6	2411	9	AF262962	AF262962 Homo sapi
3	1900.6	87.5	2189	9	HSM802178	AL137462 Homo sapi
4	1577.6	72.6	3030	10	S75275	S75275 RVLG=vasa-1
5	1562	71.9	1984	9	AK093439	AK093439 Homo sapi
6	1443	66.4	1930	10	MUSDVH	D14859 Mouse mRNA
7	751.2	34.6	1989	5	AB004836	AB004836 Gallus ga
8	712	32.8	2502	5	AF046043	AF046043 Xenopus l
9	686.8	31.6	2865	5	AB005147	AB005147 Danio rer
10	686.4	31.6	2492	5	DRY12007	Y12007 D.rerio vlg
11	619.4	28.5	2058	5	AB051835	AB051835 Oreochrom
12	619.4	28.5	2130	5	AB032467	AB032467 Oreochrom
13	612.2	28.2	2612	5	AB032566	AB032566 Oncorhync
14	565	26.0	1512	5	AF479823	AF479823 Pantodon
15	552.8	25.5	2187	5	AB063484	AB063484 Oryzias l
16	548.2	25.2	1198	5	AF251800	AF251800 Danio dan
17	543.8	25.0	1730	5	AF479821	AF479821 Hyphessob
18	511.8	23.6	1972	5	AF479820	AF479820 Cyprinus
19	503.4	23.2	1506	5	AF479824	AF479824 Melanotae
20	499.6	23.0	1586	5	AF520608	AF520608 Sparus au
21	498.6	23.0	1816	5	AF479825	AF479825 Oncorhync
22	487.4	22.4	2551	3	AB016603	AB016603 Ciona int
23	487.4	22.4	2561	3	AB016604	AB016604 Ciona int
24	457	21.0	1873	3	AB047385	AB047385 Ephydatia
25	456.6	21.0	3132	3	AB047383	AB047383 Hydra mag
26	438	20.2	3023	3	AB047803	AB047803 Ciona sav
27	436.4	20.1	2777	3	AB047802	AB047802 Ciona sav
28	434.6	20.0	2721	3	AB047382	AB047382 Hydra mag
29	428.8	19.7	1449	5	AF479822	AF479822 Oryzias l
30	383.2	17.6	2319	9	HSAF000985	AF000985 Homo sapi
31	383.2	17.6	4416	9	HSAF000984	AF000984 Homo sapi
32	378.2	17.4	2384	3	AB047381	AB047381 Hydra mag
33	377	17.4	524	4	AY100475	AY100475 Equus cab
34	371	17.1	2187	6	AX305316	AX305316 Sequence
35	371	17.1	2187	10	MUSRNAHELI	L25126 Mus musculu
36	371	17.1	2201	9	AF061337	AF061337 Homo sapi
37	371	17.1	2508	9	BC011819	BC011819 Homo sapi
38	371	17.1	3158	9	HSU50553	U50553 Homo sapien
39	371	17.1	3185	10	MMDBRNAHL	Z38117 M.musculus
40	370.2	17.0	3731	10	MMU7376	AJ007376 Mus muscu
41	370.2	17.0	3767	10	BC021453	BC021453 Mus muscu
42	369.4	17.0	1989	10	AB080116	AB080116 Mesocrice
43	369.4	17.0	3408	9	HSAF000983	AF000983 Homo sapi
44	369.4	17.0	5322	9	HSAF000982	AF000982 Homo sapi
45	358	16.5	1983	10	MUSPL10A	J04847 Mouse PL10

ALIGNMENTS

```
RESULT 1
AY004154
LOCUS       AY004154               2224 bp    mRNA    linear   PRI 25-AUG-2000
DEFINITION  Homo sapiens DEAD box RNA helicase (VASA) mRNA, complete cds.
ACCESSION   AY004154
VERSION     AY004154.1  GI:9438226
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2224)
  AUTHORS   Castrillon,D.H., Quade,B.J., Wang,T.Y., Quigley,C. and Crum,C.P.
  TITLE     The human VASA gene is specifically expressed in the germ cell
            lineage
```

```
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9585-9590 (2000)
  MEDLINE   20402578
   PUBMED   10920202
REFERENCE   2  (bases 1 to 2224)
  AUTHORS   Castrillon,D.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2000) Pathology, Brigham and Women's Hospital, 75
            Francis Street, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..2224
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q11.2-12"
                     /tissue_type="testis"
                     /dev_stage="adult"
     gene            1..2224
                     /gene="VASA"
     CDS             14..2188
                     /gene="VASA"
                     /codon_start=1
                     /product="DEAD box RNA helicase"
                     /protein_id="AAF86585.1"
                     /db_xref="GI:9438227"
                     /translation="MGDEDWEAEINPHMSSYVPIFEKDRYSGENGDNFNRTPASSSEM
                     DDGPSRRDHFMKSGFASGRNFGNRDAGECNKRDNTSTMGGFGVGKSFGNRGFSNSRFE
                     DGDSSGFWRESSNDCEDNPTRNRGFSKRGGYRDGNNSEASGPYRRGGRGSFRGCRGGF
                     GLGSPNNDLDPDECMQRTGGLFGSRRPVLSGTGNGDTSQSRSGSGSERGGYKGLNEEV
                     ITGSGKNSWKSEAEGGESSDTQGPKVTYIPPPPPEDEDSIFAHYQTGINFDKYDTILV
                     EVSGHDAPPAILTFEEANLCQTLNNNIAKAGYTKLTPVQKYSIPIILAGRDLMACAQT
                     GSGKTAAFLLPILAHMMHDGITASRFKELQEPECIIVAPTRELVNQIYLEARKFSFGT
                     CVRAVVIYGGTQLGHSIRQIVQGCNILCATPGRLMDIIGKEKIGLKQIKYLVLDEADR
                     MLDMGFGPEMKKLISCPGMPSKEQRQTLMFSATFPEEIQRLAAEFLKSNYLFVAVGQV
                     GGACRDVQQTVLQVGQFSKREKLVEILRNIGDERTMVFVETKKKADFTATFLCQEKIS
                     TTSIHGDREQREREQALGDFRFGKCPVLVATSVAARGLDIENVQHVINFDLPSTIDEY
                     VHRIGRTGRCGNTGRAISFFDLESDNHLAQPLVKVLTDAQQDVPAWLEEIAFSTYIPG
                     FSGSTRGNVFASVDTRKGKSTLNTAGFSSSRAPNPVDDESWD"
BASE COUNT      678 a    385 c    550 g    611 t
ORIGIN

  Query Match            100.0%;  Score 2172;  DB 9;  Length 2224;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2172;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 73

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 133

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 193

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 253

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 313

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 373

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 433

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 493

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 553

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 613

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 673

Qy        661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 733

Qy        721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 793

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 853

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 913

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 973

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1033

Qy       1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1093

Qy       1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1153

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1213

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1273

Qy       1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1333

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1393

Qy       1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1453

Qy       1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1513

Qy       1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1514 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1573

Qy       1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1574 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1633

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1634 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1693

Qy       1681 TGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1694 TGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1753

Qy       1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1754 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1813

Qy       1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1814 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1873

Qy       1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1874 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1933

Qy       1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1934 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1993

Qy       1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1994 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2053

Qy       2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2054 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2113

Qy       2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2114 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2173

Qy       2161 GAGTCATGGGAT 2172
              ||||||||||||
Db       2174 GAGTCATGGGAT 2185
```

```
RESULT 2
AF262962
LOCUS       AF262962                2411 bp    mRNA    linear   PRI 30-MAY-2000
DEFINITION  Homo sapiens VASA protein mRNA, complete cds.
ACCESSION   AF262962
VERSION     AF262962.1  GI:8102020
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2411)
  AUTHORS   Rocha,D. and Affara,N.
  TITLE     Cloning and characterization of the human VASA gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2411)
  AUTHORS   Rocha,D. and Affara,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAY-2000) Deparment of Pathology, University of
            Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK
FEATURES             Location/Qualifiers
     source          1..2411
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
     CDS             18..2192
                     /codon_start=1
                     /product="VASA protein"
                     /protein_id="AAF72705.1"
                     /db_xref="GI:8102021"
                     /translation="MGDEDWEAEINPHMSSYVPIFEKDRYSGENGDNFNRTPASSSEM
                     DDGPSRRDHFMKSGFASGRNFGNRDAGECNKRDNTSTMGGFGVGKSFGNRGFSNSRFE
                     DGDSSGFWRESSNDCEDNPTRNRGFSKRGGYRDGNNSEASGPYRRGGRGSFRGCRGGF
                     GLGSPNNDLDPDECMQRTGGLFGSRRPVLSGTGNGDTSQSRSGSGSERGGYKGLNEEV
                     ITGSGKNSWKSEAEGGESSDTQGPKVTYIPPPPPEDEDSIFAHYQTGINFDKYDTILV
                     EVSGHDAPPAILTFEEANLCQTLNNNIAKAGYTKLTPVQKYSIPIILAGRDLMACAQT
                     GSGKTAAFLLPILAHMMHDGITASRFKELQEPECIIVAPTRELVNQIYLEARKFSFGT
                     CVRAVVIYGGTQLGHSIRQIVQGCNILCATPGRLMDIIGKEKIGLKQIKYLVLDEADR
                     MLDMGFGPEMKKLISCPGMPSKEQRQTLMFSATFPEEIQRLAAEFLKSNYLFVAVGQV
                     GGACRDVQQTVLQVGQFSKREKLVEILRNIGDERTMVFVETKKKADFIATFLCQEKIS
                     TTSIHGDREQREREQALGDFRFGKCPVLVATSVAARGLDIENVQHVINFDLPSTIDEY
                     VHRIGRTGRCGNTGRAISFFDLESDNHLAQPLVKVLTDAQQDVPAWLEEIAFSTYIPG
                     FSGSTRGNVFASVDTRKGKSTLNTAGFSSSQAPNPVDDESWD"
BASE COUNT      771 a    408 c    571 g    661 t
ORIGIN

  Query Match             99.6%;  Score 2164;  DB 9;  Length 2411;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2167;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 77

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 137

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 197

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 257

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 317

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 377

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 437

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 497

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 557

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 617

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 677

Qy        661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 737

Qy        721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           738 GGACCAAAAGTGACCTACATACCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 797

Qy           781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           798 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 857

Qy           841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           858 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 917

Qy           901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           918 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 977

Qy           961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
                 |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db           978 ATCATACTTGCAGGACGAGATTTGATGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGCG 1037

Qy          1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1038 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1097

Qy          1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1098 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1157

Qy          1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1158 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1217

Qy          1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1218 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1277

Qy          1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1278 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1337

Qy          1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1338 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1397

Qy          1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1398 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1457

Qy          1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1458 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1517

Qy          1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1518 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1577

Qy          1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1578 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1637

Qy          1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
                 |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db          1638 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTATTGCAACTTTTCTT 1697

Qy          1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
                 ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1698 TGTCAAGAAAAAATATCAACTACAAGTATTCATGGTGATCGGGAACAGAGAGAGCGGGAG 1757

Qy          1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1758 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1817

Qy          1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1818 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1877

Qy          1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1878 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1937

Qy          1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1938 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1997

Qy          1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1998 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2057

Qy          2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2058 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2117

Qy          2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
                 ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db          2118 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACAAGCTCCCAATCCAGTAGATGAT 2177

Qy          2161 GAGTCATGGGAT 2172
                 ||||||||||||
Db          2178 GAGTCATGGGAT 2189
```

```
RESULT 3
HSM802178
LOCUS       HSM802178               2189 bp    mRNA    linear   PRI 18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp434B1122 (from clone DKFZp434B1122);
            partial cds.
ACCESSION   AL137462
VERSION     AL137462.1  GI:6808043
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2189)
  AUTHORS   Blum,H., Bauersachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by LMU (Ludwig Maximilians University,
            Munich/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp434B1122) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1..2189
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434B1122"
                     /tissue_type="testis"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     gene            1..1910
                     /gene="DKFZp434B1122"
     CDS             <1..1910
                     /gene="DKFZp434B1122"
                     /note="strong similarity to testis specific Helicase (VASA
                     homolog)"
                     /codon_start=3
                     /product="hypothetical protein"
                     /protein_id="CAB70750.1"
                     /db_xref="GI:6808044"
```

```
                    /translation="SFGNRGFSNSRFEDGDSSGFWRESSNDCEDNPTRNRGFSKRGGY
                    RDGNNSEASGPYRRGGRGSFRGCRGGFGLGSPNNDLDPDECMQRTGGLFGSRRPVLSG
                    TGNGDTSQSRSGSGSERGGYKGLNEEVITGSGKNSWKSEAEGGESSDTQGPKVTYIPP
                    PPPEDEDSIFAHYQTGINFDKYDTILVEVSGHDAPPAILTFEEANLCQTLNNNIAKAG
                    YTKLTPVQKYSIPIILAGRDLMACAQTGSGKTAAFLLPILAHMMHDGITASRFKELQE
                    PECIIVAPTRELVNQIYLEARKFSFGTCVRAVVIYGGTQLGHSIRQIVQGCNILCATP
                    GRLMDIIGKEKIGLKQIKYLVLDEADRMLDMGFGPEMKKLISCPGMPSKEQRQTLMFS
                    ATFPEEIQRLAAEFLKSNYLFVAVGQVGGACRDVQQTVLQVGQFSKREKLVEILRNIG
                    DERTMVFVETKKKADFIATFLCQEKISTTSIHGDREQREREQALGDFRFGKCPVLVAT
                    SVAARGLDIENVQHVINFDLPSTIDEYVHRIGRTGRCGNTGRAISFFDLESDNHLAQP
                    LVKVLTDAQQDVPAWLEEIAFSTYIPGFSGSTRGNVFASVDTRKGKSTLNTAGFSSSQ
                    APNPVDDESWD"
     polyA_signal   2128..2133
     polyA_site     2147
BASE COUNT      706 a    373 c    505 g    605 t
ORIGIN

  Query Match             87.5%;  Score 1900.6;  DB 9;  Length 2189;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1903;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        266 AGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTT 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTT 60

Qy        326 TCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCA 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCA 120

Qy        386 AGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTG 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTG 180

Qy        446 GAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACT 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACT 240

Qy        506 TAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTAT 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTAT 300

Qy        566 TAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAG 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAG 360

Qy        626 GTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGT 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGT 420

Qy        686 CAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCC 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCC 480

Qy        746 CTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCG 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCG 540

Qy        806 ACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGA 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGA 600

Qy        866 CTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATA 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATA 660

Qy        926 CTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGA 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGA 720

Qy        986 TGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTC 1045
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTC 780

Qy       1046 ATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTA 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTA 840

Qy       1106 TTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTT 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTT 900

Qy       1166 CTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAA 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAA 960

Qy       1226 TTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATA 1285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATA 1020

Qy       1286 TCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTG 1345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTG 1080

Qy       1346 ATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAA 1405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAA 1140

Qy       1406 TGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTC 1465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTC 1200

Qy       1466 AAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGG 1525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGG 1260

Qy       1526 GTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAG 1585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAG 1320

Qy       1586 AAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAA 1645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAA 1380

Qy       1646 CTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTGTCAAGAAAAAATATCAACTACAA 1705
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1381 CTAAGAAAAAAGCAGATTTTATTGCAACTTTTCTTGTCAAGAAAAAATATCAACTACAA 1440

Qy       1706 GTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTG 1765
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTATTCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTG 1500

Qy       1766 GAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATG 1825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATG 1560

Qy       1826 TGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTG 1885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTG 1620

Qy       1886 GGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGG 1945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGG 1680

Qy       1946 ATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTG 2005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTG 1740

Qy       2006 CATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAG 2065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAG 1800

Qy       2066 GAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGT 2125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGT 1860

Qy       2126 TTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
```

```
Db       1861 TTTCTTCTTCACAAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 1907

RESULT 4
S75275
LOCUS       S75275                  3030 bp    mRNA    linear   ROD 15-MAY-1995
DEFINITION  RVLG=vasa-like gene protein [rats, Wistar-Imanishi, testis, mRNA,
            3030 nt].
ACCESSION   S75275
VERSION     S75275.1  GI:806463
KEYWORDS    .
SOURCE      Rattus sp. testis Wistar-Imanishi.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3030)
  AUTHORS   Komiya,T. and Tanigawa,Y.
  TITLE     Cloning of a gene of the DEAD box protein family which is
            specifically expressed in germ cells in rats
  JOURNAL   Biochem. Biophys. Res. Commun. 207 (1), 405-410 (1995)
  MEDLINE   95160706
   PUBMED   7857296
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 159937] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1..3030
                     /organism="Rattus sp."
                     /db_xref="taxon:10118"
     gene            1..3030
                     /gene="RVLG"
                     /note="vasa-like gene protein, RVLG protein"
     CDS             266..2407
                     /gene="RVLG"
                     /note="putative DEAD box protein; This sequence comes from
                     Fig. 1; RVLG protein"
                     /codon_start=1
                     /product="vasa-like gene protein"
                     /protein_id="AAB33364.1"
                     /db_xref="GI:806464"
                     /translation="MGDEDWEAEILKPHVSSYVPVFEKDKYSSGANGDTFNRTSASSS
                     EMEDGPSGRDHFMRSGFSSGRNLGNRDIGESSKRETTSTTGGFGRGKGFGNRGFLNNK
                     FEEGDSSGFWKESTNDCEDTQTRSRGFSKRGGYPDGNDSEASGPFRRGGRDSEYDQDQ
                     GSQRGGGLFGSRKPAASDSGSGDTFQSRSGNARGAYKGLNEEVVTGSGKNSWKSEAEG
                     GESSDIQGPKVTYIPPPPPEDEDSIFAHYQTGINFDKYDTILVEVSGHDAPPAILTFE
                     EANLCQTLNNNIAKAGYTKLTPVQKYSIPIVLAGRDLMACAQTGSGKTAAFLLPILAH
                     MMRDGITASRFKELQEPECIIVAPTRELINQIYLEARKFSFGTCVRAVVIYGGTQFGH
                     SIRQIVQGCNILCATPGRLMDIIGKEKIGLKQVKYLVLDEADRMLDMGFGPEMKKLIS
                     CPGMPSKEQRQTLLFSATFPEEIQRLAGEFLKSNYLFVAVGQVGGACRDVQQSILQVG
                     PVFKKRKLVEILRNIGDERPMVFVETKKKADFIATFLCQEKISTTSIHGDREQREREQ
                     ALGDFRCGKCPVLVATSVAARGLDIENVQHVINFNLPSTIDEYVHRIGRTGRCGNTGR
                     AISFFDTESDNHLAQPLVKVLSDAQQDVPAWLEEIAFSSYAPPSFSNSTRGAVFASFD
                     TRKNFQGKNTLNTAGISSAQAPNPVDDESWD"
BASE COUNT      900 a    551 c    723 g    856 t
ORIGIN

  Query Match             72.6%;  Score 1577.6;  DB 10;  Length 3030;
  Best Local Similarity   85.0%;  Pred. No. 0;
  Matches 1862;  Conservative    0;  Mismatches  259;  Indels   69;  Gaps    6;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAA---TCAACCCTCATATGTCTTCCTATGTTCCC 57

Db        266 ATGGGAGATGAAGATTGGGAGGCAGAAATCCTCAAACCTCATGTGTCTTCCTACGTTCCT 325

Qy         58 ATATTTGAGAAGGATAGGTA---TTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCA 114

Db        326 GTATTTGAGAAGGATAAATATTCTTCTGGGGCAAATGGAGACACTTTTAACAGGACTTCA 385

Qy        115 GCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGA 174

Db        386 GCTTCATCATCAGAAATGGAAGATGGACCTTCTGGAAGAGATCATTTCATGAGAAGTGGA 445

Qy        175 TTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT 234

Db        446 TTTTCCTCTGGAAGAAATTTAGGAAACAGAGATATTGGCGAGTCTAGTAAAAGAGAGACT 505

Qy        235 ACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAAC 294

Db        506 ACATCTACAACCGGTGGCTTTGGAAGAGGAAAGGGTTTTGGAAACAGAGGTTTTTTAAAT 565

Qy        295 AGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAA 354

Db        566 AACAAGTTTGAAGAAGGTGACAGCTCTGGTTTCTGGAAAGAGTCTACTAATGACTGTGAA 625

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414

Db        626 GATACTCAGACTCGGAGCAGAGGGTTTTCCAAGCGAGGCGGCTATCCAGATGGGAATGAT 685

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGA 474

Db        686 TCGGAAGCTTCAGGCCCATTCAGAAGAGGTGGGAGAG---------------------- 722

Qy        475 GGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACT 534

Db        723 ----------------------ATAGTGAATATGACCAAGATCAGGGATCACAGCGTGGT 760

Qy        535 GGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCT 594

Db        761 GGTGGCCTTTTTGGTTCTAGGAAACCAGCAGCAAGTGATTCAGGCAGTGGTGACACTTTC 820

Qy        595 CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654

Db        821 CAGAGCAGA------AGTGGGAATGCCCGAGGTGCCTACAAAGGCTTAAATGAAGAAGTA 874

Qy        655 ATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGAT 714

Db        875 GTAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCTGAAGGAGGCGAAAGCAGTGAT 934

Qy        715 ACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATC 774

Db        935 ATTCAAGGTCCAAAAGTGACATATATACCCCCTCCTCCACCAGAGGATGAGGACTCCATC 994

Qy        775 TTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTG 834

Db        995 TTTGCACATTATCAGACAGGCATAAACTTTGATAAATATGATACCATACTTGTTGAAGTA 1054

Qy        835 TCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACA 894

Db       1055 TCTGGACATGATGCACCACCGGCAATTTTGACTTTTGAAGAAGCGAATCTCTGCCAGACC 1114

Qy        895 CTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGT 954

Db       1115 CTGAATAACAACATTGCTAAGGCTGGCTATACCAAGCTCACTCCTGTGCAGAAGTACAGC 1174

Qy        955 ATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAG 1014

Db       1175 ATTCCCATTGTGTTAGCAGGAAGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAG 1234

Qy       1015 ACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGT 1074

Db       1235 ACGGCAGCTTTTCTCTTGCCTATTTTGGCTCATATGATGAGGGATGGAATAACTGCCAGT 1294

Qy       1075 CGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTC 1134

Db       1295 CGCTTTAAAGAACTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGATC 1354

Qy       1135 AACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTT 1194

Db       1355 AACCAAATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTC 1414

Qy       1195 ATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATA 1254

Db       1415 ATATATGGAGGAACCCAGTTTGGGCACTCAATTCGACAGATAGTGCAAGGGTGTAATATA 1474
```

```
Qy       1255 TTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAA 1314
Db       1475 TTATGTGCTACTCCAGGGAGGCTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAA 1534

Qy       1315 CAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCA 1374
Db       1535 CAAGTCAAGTACTTAGTTTTGGATGAAGCTGATCGAATGTTGGATATGGGTTTTGGACCC 1594

Qy       1375 GAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTT 1434
Db       1595 GAAATGAAGAAGTTAATTTCTTGTCCAGGAATGCCATCAAAGGAACAGCGCCAAACTCTC 1654

Qy       1435 ATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCA 1494
Db       1655 TTATTCAGTGCAACTTTTCCAGAAGAAATCCAGAGGTTGGCTGGCGAGTTTTTAAAGTCA 1714

Qy       1495 AATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACC 1554
Db       1715 AATTATTTGTTTGTTGCTGTTGGACAAGTGGGAGGAGCTTGCAGAGATGTGCAGCAGTCC 1774

Qy       1555 GTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATA 1614
Db       1775 ATTCTTCAAGTTGGCCCAGTATTCAAAAAGAGAAAACTTGTTGAGATTCTACGAAACATA 1834

Qy       1615 GGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
Db       1835 GGTGATGAAAGACCTATGGTCTTTGTTGAAACCAAGAAAAAAGCAGATTTCATTGCGACT 1894

Qy       1675 TTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
Db       1895 TTCCTTTGTCAAGAAAAAATATCAACTACAAGTATTCATGGCGATCGGGAACAGAGGGAG 1954

Qy       1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
Db       1955 CGAGAACAAGCTCTTGGAGATTTTCGCTGTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 2014

Qy       1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
Db       2015 GTGGCTGCCAGAGGACTTGATATTGAAAATGTTCAACATGTTATCAATTTTAACCTTCCT 2074

Qy       1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
Db       2075 TCTACCATTGATGAATATGTTCATCGAATTGGACGTACTGGACGTTGTGGAAATACTGGC 2134

Qy       1915 AGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAA 1974
Db       2135 AGGGCGATTTCTTTTTTTGATACCGAATCTGATAATCATTTAGCACAACCTCTAGTTAAA 2194

Qy       1975 GTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACA 2034
Db       2195 GTACTGTCAGATGCTCAACAGGATGTTCCTGCGTGGTTAGAAGAGATTGCCTTTAGTTCA 2254

Qy       2035 TA---CATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACC 2091
Db       2255 TATGCGCCTCCCAGCTTCAGTAATAGCACAAGAGGGGCTGTGTTTGCATCTTTTGACACT 2314

Qy       2092 AGAA---------AGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
Db       2315 AGGAAGAATTTCCAGGGCAAGAACACACTGAACACAGCTGGGATTTCTTCTGCACAAGCT 2374

Qy       2143 CCCAATCCAGTAGATGATGAGTCATGGGAT 2172
Db       2375 CCCAATCCAGTTGATGACGAGTCATGGGAT 2404
```

```
RESULT 5
AK093439
LOCUS       AK093439                1984 bp    mRNA    linear   PRI 15-JUL-2002
DEFINITION  Homo sapiens cDNA FLJ36120 fis, clone TESTI2022738, highly similar
            to DEAD BOX PROTEIN 4.
ACCESSION   AK093439
VERSION     AK093439.1  GI:21752310
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens testis cDNA to mRNA, clone_lib:TESTI2
            clone:TESTI2022738.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1984)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1..1984
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TESTI2022738"
                     /tissue_type="testis"
                     /clone_lib="TESTI2"
                     /note="cloning vector: pME18SFL3"
BASE COUNT      612 a    357 c    448 g    567 t
ORIGIN

  Query Match             71.9%;  Score 1562;  DB 9;  Length 1984;
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches 1623;  Conservative    0;  Mismatches    5;  Indels   48;  Gaps    1;

Qy        497 ATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAA 556
Db        117 ATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAA 176

Qy        557 GACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAA 616
Db        177 GACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAA 236

Qy        617 GTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATT 676
Db        237 GTGAACGAG-----------------------------------------------ATT 248

Qy        677 CTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCT 736
Db        249 CTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCGAGGACCAAAAGTGACCT 308

Qy        737 ACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCA 796
Db        309 ACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCA 368

Qy        797 TAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAG 856
Db        369 TAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAG 428

Qy        857 CAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAG 916
Db        429 CAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAG 488

Qy        917 CTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGAC 976
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 CTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGAC 548

Qy        977 GAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAA 1036
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        549 GAGATTTGATGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAA 608

Qy       1037 TTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAAC 1096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 TTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAAC 668

Qy       1097 CAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCA 1156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 CAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCA 728

Qy       1157 GAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGG 1216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGG 788

Qy       1217 GACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGAC 1276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 GACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGAC 848

Qy       1277 TGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGG 1336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 TGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGG 908

Qy       1337 ATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTT 1396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 ATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTT 968

Qy       1397 GCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAG 1456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 GCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAG 1028

Qy       1457 AGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTG 1516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 AGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTG 1088

Qy       1517 GACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCT 1576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 GACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCT 1148

Qy       1577 CAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCT 1636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1149 CAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCT 1208

Qy       1637 TTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATAT 1696
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1209 TTGTTGAAACTAAGAAAAAAGCAGATTTTATTGCAACTTTTCTTTGTCAAGAAAAAAATAT 1268

Qy       1697 CAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATT 1756
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 CAACTACAAGTATTCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATT 1328

Qy       1757 TTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATA 1816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 TTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATA 1388

Qy       1817 TTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTC 1876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 TTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTC 1448

Qy       1877 ATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATC 1936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 ATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATC 1508

Qy       1937 TTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGG 1996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGG 1568

Qy       1997 ATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTA 2056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTA 1628

Qy       2057 GTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACA 2116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1629 GTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACA 1688

Qy       2117 CAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1689 CAGCTGGGTTTTCTTCTTCACAAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 1744
```

```
RESULT 6
MUSDVH
LOCUS       MUSDVH                  1930 bp    mRNA    linear   ROD 04-FEB-1999
DEFINITION  Mouse mRNA for drosophila vasa homologue, partial cds.
ACCESSION   D14859
VERSION     D14859.1  GI:286074
KEYWORDS    RNA helicase; drosophila vasa homologue.
SOURCE      Mus musculus (strain BALB/c) (library: lambda gt10) adult and
            embryo gonad and testis primordial germ cell, spermatogonium and
            spermatocyte cDNA to mRNA, clone mVH.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (sites)
  AUTHORS   Fujiwara,Y., Komiya,T., Kawabata,H., Sato,M., Fujimoto,H.,
            Furusawa,M. and Noce,T.
  TITLE     Isolation of a DEAD-family protein gene that encodes a murine
            homolog of Drosophila vasa and its specific expression in germ cell
            lineage
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 91 (25), 12258-12262 (1994)
  MEDLINE   95083681
REFERENCE   2  (bases 1 to 1930)
  AUTHORS   Noce,T.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1930)
  AUTHORS   Noce,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-1993) Toshiaki Noce, Mitsubishi Kasei Institute
            of Life Sciences, Developmental Biology; 11 Minamiooya, Machida,
            Tokyo 194, Japan (Tel:0427-24-6246, Fax:0427-29-1252)
FEATURES             Location/Qualifiers
     source          1..1930
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /cell_type="primordial germ cell, spermatogonium,
                     spermatocyte"
                     /tissue_type="gonad, testis"
                     /clone_lib="lambda gt10"
                     /dev_stage="adult, embryo"
     CDS             <1..1914
                     /function="RNA helicase"
                     /codon_start=1
                     /product="Drosophila vasa homologue"
                     /protein_id="BAA03584.1"
                     /db_xref="GI:286075"
                     /translation="FGRGKGFGNRGFLNNKFEEGDSSGFWKESNNDCEDNQTRSRGFS
                     KRGGCQDGNDSEASGPFRRGGRGSFRGCRGGFGLGRPNSESDQDQGTQCGGGFLVLGK
                     PAASDSGNGDTYQSRSGSGRGGYKGLNEEVVTGSGKNSWKSETEGGESSDSQGPKVTY
                     IPPPPPEDEDSIFAHYQTGINFDKYDTILVEVSGHDAPPAILTFEEANLCQTLNNNIR
                     KAGYTKLTPVQKYTIPIVLAGRDLMACAQTGSGKTAAFLLPILAHMMRDGITASRFKE
                     LQEPECIIVAPTRELINQIYLEARKFSFGTCVISVVIYGGTQFGHSVRQIVQGCNILC
                     ATPGRLMDIIGKEKIGLKQVKYLVLDEADSMLDMGFAPEIKKLISCPGMPSKEQHQTL
                     LFSATFPEEIQRLAGDFLKSNYLFVAVGQVGGACRDVQQTILQVGQYQKEKSLLRFYE
                     NIGDERTMVFVETKKKADFIATFLCQEKISSTSIHGDREQREREQALGDFRCGKCPVL
                     VATSVAARGLDIENVQHVINFDLPSTIDEYVHRIGRTGRCGNTGRAISFFDTDSDNHL
                     AQPLVKVLSDAQQDVPAWLEEIAFSTYVPPSFSSSTRGGAVFASVDTRKNYQGKAHVE
                     YSGDFFFTSSQSS"
BASE COUNT      577 a    356 c    487 g    510 t
ORIGIN
```

```
Query Match             66.4%;  Score 1443;  DB 10;  Length 1930;
Best Local Similarity   86.4%;  Pred. No. 0;
Matches 1671;  Conservative    0;  Mismatches  240;  Indels   24;  Gaps    6;

Qy        253 TTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGT 312
Db          1 TTTGGAAGAGGAAAGGGCTTTGGAAACAGAGGTTTTTAAATAACAAGTTTGAAGAAGGT 60

Qy        313 GATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAAC 372
Db         61 GATAGCTCTGGTTTCTGGAAAGAGTCTAATAATGACTGTGAAGATAATCAGACTCGAAGC 120

Qy        373 AGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCA 432
Db        121 AGAGGGTTTTCCAAGCGAGGTGGCTGCCAAGATGGAAACGATTCAGAAGCATCAGGCCCG 180

Qy        433 TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGT 492
Db        181 TTCAGAAGAGGGGGAAGAGGCAGTTTCCGAGGCTGCCGTGGAGGATTTGGTCTAGGAAGA 240

Qy        493 CCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCT 552
Db        241 CCAAATAGTGAATCTGACCAAGATCAGGGGACACAGTGTGGTGGTGGCTTTCTGGTTCTA 300

Qy        553 AGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGT 612
Db        301 GGGAAACCAGCAGCAAGTGATTCAGGCAATGGTGACACTTACCAAAGCAGA------AGT 354

Qy        613 GGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAG 672
Db        355 GGAAGTGGTCGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAGTAACAGGTTCTGGAAAG 414

Qy        673 AATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTG 732
Db        415 AATTCTTGGAAGTCAGAAACTGAAGGAGGTGAAAGCAGTGATAGTCAAGGTCCAAAAGTG 474

Qy        733 ACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACA 792
Db        475 ACATATATACCCCCCCCTCCACCAGAGGATGAGGACTCCATCTTTGCACATTATCAGACA 534

Qy        793 GGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCA 852
Db        535 GGCATAAACTTTGATAAATATGACACCATACTTGTTGAAGTATCTGGACATGATGCACCA 594

Qy        853 CCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCT 912
Db        595 CCGGCAATTTTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTCGT 654

Qy        913 AAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCA 972
Db        655 AAAGCTGGCTATACTAAGCTTACTCCTGTGCAGAAGTACACGATTCCCATTGTATTAGCA 714

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
Db        715 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCAGCTTTTCTCTTG 774

Qy       1033 CCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAG 1092
Db        775 CCTATTTTGGCTCATATGATGCGGGATGGAATAACTGCCAGTCGCTTTAAAGAACTGCAG 834

Qy       1093 GAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAA 1152
Db        835 GAACCAGAGTGTATTATTGTAGCACCAACTCGAGAACTGATCAACCAAATTTACTTGGAA 894

Qy       1153 GCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAG 1212
Db        895 GCCAGAAAATTTTCTTTTGGGACTTGTGTAATATCTGTGGTCATATATGGAGGGACCCAG 954

Qy       1213 CTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGA 1272
Db        955 TTTGGTCATTCAGTTCGACAGATAGTACAAGGGTGTAACATATTGTGTGCTACTCCAGGG 1014

Qy       1273 AGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTT 1332
Db       1015 AGGCTGATGGACATCATAGGCAAAGAAAAGATTGGCCTCAAACAAGTCAAGTACTTAGTT 1074

Qy       1333 TTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATT 1392
Db       1075 TTGGATGAAGCTGATAGCATGCTAGACATGGGTTTTGCACCAGAGATCAAGAAACTAATT 1134

Qy       1393 TCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTT 1452
Db       1135 TCTTGTCCAGGAATGCCATCAAAGGAACAACACCAAACCCTTTTATTCAGTGCTACTTTT 1194

Qy       1453 CCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCT 1512
Db       1195 CCAGAAGAAATCCAGAGGTTGGCTGGGGATTTTCTAAAGTCCAATTATTTGTTTGTTGCT 1254

Qy       1513 GTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAG 1572
Db       1255 GTTGGGCAAGTGGGAGGAGCTTGCAGAGATGTTCAGCAGACGATTCTTCAAGTTGGCCAG 1314

Qy       1573 TTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCG-AAACATAGGGGATGAAAGAACTAT 1631
Db       1315 -TATCAAAAAGAGAAAAGCTTGTTGAGATTCTACGAAAACATAGGTGATGAAAGAACTAT 1373

Qy       1632 GGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAA 1691
Db       1374 GGTCTTTGTTGAAACCAAGAAAAAAGCCGATTTCATTGCAACTTTTCTTTGTCAAGAAAA 1433

Qy       1692 AATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGG 1751
Db       1434 AATATCAAGTACAAGTATTCATGGTGATCGGGAACAGAGGGAGAGAGAGCAAGCTCTTGG 1493

Qy       1752 AGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCT 1811
Db       1494 AGATTTTCGCTGTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTGGCTGCCAGAGGGCT 1553

Qy       1812 GGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATA 1871
Db       1554 TGATATTGAAAATGTTCAACATGTTATCAATTTTGACCTTCCTTCTACCATTGATGAGTA 1613

Qy       1872 TGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTT 1931
Db       1614 TGTTCATCGAATTGGACGCACTGGCCGCTGTGGAAATACTGGCAGAGCGATTTCTTTTTT 1673

Qy       1932 TGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCA 1991
Db       1674 TGATACTGACTCTGATAATCATTTAGCACAGCCTCTAGTTAAAGTACTGTCAGACGCTCA 1733

Qy       1992 ACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACAT---TCCTGGCTT 2048
Db       1734 ACAGGATGTCCCCGCATGGCTAGAAGAGATTGCCTTCAGTACCTATGTGCCTCCCAGCTT 1793

Qy       2049 CAGTGGTAGTACAAGA---GGAAACGTGTTTGCATCAGTTGATACCAGAAA--------- 2096
Db       1794 CAGTAGCAGCACAAGAGGGGGGGCCGTGTTTGCATCTGTTGACACGAGGAAGAATTACCA 1853

Qy       2097 -GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAG 2155
Db       1854 GGGCAAAGCACACGTTGAATACAGCGGGGATTTCTTCTTCACAAGCTCCCAATCCAGTTG 1913

Qy       2156 ATGATGAGTCATGGG 2170
Db       1914 ATGACGAATCATGGG 1928
```

```
RESULT 7
AB004836
LOCUS       AB004836                1989 bp    mRNA    linear   VRT 01-SEP-2000
DEFINITION  Gallus gallus mRNA for Cvh, complete cds.
ACCESSION   AB004836
VERSION     AB004836.1  GI:9967267
KEYWORDS    Cvh.
SOURCE      Gallus gallus cDNA to mRNA.
  ORGANISM  Gallus gallus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (sites)
  AUTHORS   Tsunekawa,N., Naito,M., Sakai,Y., Nishida,T. and Noce,T.
  TITLE     Isolation of chicken vasa homolog gene and tracing the origin of
            primordial germ cells
  JOURNAL   Development 127 (12), 2741-2750 (2000)
  MEDLINE   20283536
REFERENCE   2  (sites)
  AUTHORS   Tsunekawa,N., Fujimoto,H., Nishida,T. and Noce,T.
  TITLE     Isolation of a DEAD-family protein gene that encodes a chicken
            homolog of Drosophila vasa and its specific expression in germ cell
            lineage
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1989)
  AUTHORS   Tsunekawa,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUN-1997) Naoki Tsunekawa, Mitsubishi Kasei Institute
            of Life Sciences, Molecular Reproduction and Development; 11
            Minamiooya, Machida, Tokyo 194, Japan
            (E-mail:tsune@libra.ls.m-kagaku.co.jp, Tel:+81-427-24-6246,
            Fax:+81-427-24-6314)
FEATURES             Location/Qualifiers
     source          1..1989
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
     CDS             1..1989
                     /codon_start=1
                     /product="Cvh"
                     /protein_id="BAB12337.1"
                     /db_xref="GI:9967268"
                     /translation="MEEDWDTELEQEAAAASQGRSEEQAWMANSGRPNSPSLRFSSRP
                     SSPLSGFPGRPNSPFFGFSQNKGSLGANEGLNRSLPVQHDIGGYSGSRESVVRQNRED
                     QPVTRFGRGRSSGSRDFQERNSANDPGMQDQGFRRVPGIFGQSKCFNSEERNSPLRGS
                     PFAPGGRGAVGGPAGVLKGRSEEIDSGRGPKVTYVPPPPPEDEQSIFACYQSGINFDK
                     YDECAVEMSGLDPPAPLLAFEEANFAQTLRKNISKTGYSKLTPVQKHSIPVIQAGRDL
                     MSCAQTGSGKTAAFLLPIVDRMMKDGVTASFPKQQDPQCIIVAPTRELINQIFLEARK
                     FVYGTCIRPVVIYGGTQTGHSIRQIMQGCNILCATPGRLLDIIEKGKISLVEVKYLVL
                     DEADRMLDMGFGLDMKKLISYPEMPSKDRRQTLMFSATFPEEVQRLAGEFLKTDYIFL
                     VIGNTCGACSDVQQNILQVPRLSKRDKLIEILQSTGGERTMVFVDTKKKADYLAAFLC
                     QENLPSTSIHGDREQREREIALRDFRSGKCQILVATSVASRGLDIENVQHVINFDLPN
                     TIEDYVHRIGRTGRCGNTGKAVSFFDDQSDGHLVQSLLKVLSRTQQEFQFGGRMAVQR
                     TNIVASTWCPKGLMQAVAEWNPREMRMSYSETTFKSWE"
BASE COUNT      597 a    385 c    493 g    514 t
ORIGIN

  Query Match            34.6%;  Score 751.2;  DB 5;  Length 1989;
  Best Local Similarity  71.1%;  Pred. No. 8.6e-167;
  Matches 1009;  Conservative    0;  Mismatches  408;  Indels    3;  Gaps    1;

Qy        648 AGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAG 707
Db        492 AGGAGGAAGAGGAGCAGTTGGAGGTCCTGCAGGAGTTCTCAAAGGACGCTCTGAAGAAAT 551

Qy        708 TAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGA 767
Db        552 TGATTCTGGAAGAGGTCCAAAGGTGACTTATGTCCCCCCTCCTCCACCTGAAGATGAACA 611

Qy        768 CTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGT 827
Db        612 GTCCATCTTTGCATGTTATCAGTCAGGAATTAATTTTGACAAGTATGATGAATGTGCTGT 671

Qy        828 GGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTG 887
Db        672 TGAGATGTCAGGACTTGACCCTCCAGCACCATTACTGGCTTTTGAAGAAGCTAACTTTGC 731

Qy        888 TCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAA 947
Db        732 TCAGACTTTAAGGAAGAATATATCTAAAACTGGATATTCAAAACTTACTCCAGTGCAGAA 791

Qy        948 ATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTC 1007
Db        792 GCACAGCATTCCTGTTATACAAGCAGGGCGGGATTTAATGTCATGTGCCCAGACAGGATC 851

Qy       1008 TGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAAC 1067
Db        852 AGGAAAAACAGCAGCTTTTCTTCTACCAATTGTGGACCGGATGATGAAAGATGGTGTAAC 911

Qy       1068 TGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGA 1127
Db        912 TGCAAGC---TTCCCAAAGCAGCAAGACCCACAATGCATTATTGTTGCACCAACTAGAGA 968

Qy       1128 ATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGC 1187
Db        969 ACTGATAAATCAGATCTTCTTAGAAGCAAGGAAGTTTGTGTATGGGACTTGTATAAGGCC 1028

Qy       1188 TGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTG 1247
Db       1029 TGTTGTGATCTATGGAGGTACACAGACAGGTCATTCAATCCGTCAAATAATGCAAGGCTG 1088

Qy       1248 TAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGG 1307
Db       1089 TAATATATTATGTGCCACTCCTGGAAGGCTTCTTGACATTATTGAAAAAGGGAAGATCAG 1148

Qy       1308 TCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTT 1367
Db       1149 TTTGGTGGAGGTGAAATATTTGGTACTAGATGAAGCAGACCGCATGCTCGATATGGGTTT 1208

Qy       1368 TGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCA 1427
Db       1209 TGGATTAGATATGAAGAAGCTGATTTCTTATCCAGAAATGCCATCTAAAGACAGACGTCA 1268

Qy       1428 AACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTT 1487
Db       1269 AACATTAATGTTTAGTGCCACTTTTCCTGAGGAAGTTCAAAGGCTGGCTGGTGAATTTTT 1328

Qy       1488 AAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCA 1547
Db       1329 GAAAACGGACTATATATTTCTTGTTATTGGAAATACCTGTGGAGCCTGCAGTGATGTTCA 1388

Qy       1548 GCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCG 1607
Db       1389 GCAAAATATTCTTCAGGTTCCCCGGTTATCCAAGAGGGATAAACTAATAGAAATTCTACA 1448

Qy       1608 AAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTAC 1667
Db       1449 AAGCACAGGTGGTGAACGAACCATGGTGTTTGTGGACACAAAGAAAAAAGCAGATTACCT 1508

Qy       1668 TGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACA 1727
Db       1509 TGCAGCCTTTCTTTGTCAAGAGAACCTACCATCCACCAGCATTCATGGAGATAGGGAACA 1568

Qy       1728 GAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGC 1787
Db       1569 GAGAGAGAGAGAGATAGCTCTTCGCGATTTCCGTTCTGGAAAATGTCAAATTCTTGTGGC 1628

Qy       1788 TACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGA 1847
Db       1629 AACTTCGGTAGCATCAAGAGGCCTGGATATTGAAAATGTTCAACATGTTATTAATTTTGA 1688

Qy       1848 TCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAA 1907
Db       1689 TCTCCCTAACACCATTGAAGATTATGTACATCGAATTGGACGAACTGGTCGTTGTGGAAA 1748

Qy       1908 TACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCT 1967
Db       1749 TACTGGCAAAGCAGTTTCATTCTTTGATGATCAGTCAGATGGCCATCTTGTACAATCACT 1808

Qy       1968 AGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTT 2027
Db       1809 ACTTAAAGTGCTTTCCAGAACCCAGCAGGAATTCCAGTTTGGTGGAAGAATGGCTGTCCA 1868

Qy       2028 TAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
Db       1869 AAGAACAAATATTGTTGCTTCAACTTGGTGCCCAAAGGGA 1908
```

```
RESULT 8
AF046043
LOCUS       AF046043                2502 bp    mRNA    linear   VRT 19-FEB-1998
DEFINITION  Xenopus laevis DEAD box protein (VLG1) mRNA, complete cds.
ACCESSION   AF046043 S69534
VERSION     AF046043.1  GI:2896106
KEYWORDS    .
SOURCE      Xenopus laevis.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 2502)
  AUTHORS   Komiya,T., Itoh,K., Ikenishi,K. and Furusawa,M.
  TITLE     Isolation and characterization of a novel gene of the DEAD box
            protein family which is specifically expressed in germ cells of
            Xenopus laevis
  JOURNAL   Dev. Biol. 162 (2), 354-363 (1994)
  MEDLINE   94200507
   PUBMED   8150200
REFERENCE   2  (bases 1 to 2502)
  AUTHORS   Komiya,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-1998) ERATO, JST, Hirohashi Cell Configuration
            Project, 5-9-9 Tokodai, Tsukuba 300-26, Japan
COMMENT     On Feb 19, 1998 this sequence version replaced gi:546292.
FEATURES             Location/Qualifiers
     source          1..2502
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /cell_type="germ cells"
     gene            1..2502
                     /gene="VLG1"
                     /note="similar to Drosophila vasa-like gene"
     5'UTR           1..65
                     /gene="VLG1"
     CDS             66..2168
                     /gene="VLG1"
                     /codon_start=1
                     /product="DEAD box protein"
                     /protein_id="AAC03114.1"
                     /db_xref="GI:2896107"
                     /translation="MEENWDTEIETEKPTYVPNFSTLETENTDNYSAYSNNDINNQNY
                     DSERSFGNRGGYRSERSRPSNFNRGSRTERGRGRGFGTNRNDNYSSERDVFGDDERDQ
                     RRGFPGRGGYNGNEDGQKPNAFRGRGGFRNENEQRRGFGERGGFRSENGQRNFDNRGD
                     FGNSGEEEDRPRSYGRGGFNNSDTGGRGRRGGRGGGSQYGGYKGRNEEVGVESGKSQE
                     EGNEKDEKPKKVTYIPPPPPDGEDNIFRQYQSGINFDKYDEILVDVTGKDVPPAILTF
                     EEANLCETLRRNVARAGYVKLTPVQKHSIPIIMAGRDLMACAQTGSGKTAAFLLPILS
                     YMMNEGITASQYLQLQEPEAIIIAPTRELINQIYLDARKFSYGTCVRPVVVYGGIQPV
                     HAMRDVEKGCNILCATPGRLLDIVSKEKIGLSKLRYLVLDEADRMLDMGFAPEIEKLM
                     TKPGMPTKEKRQTLMFSATYPEEIRRLASNYLKSEHLFVVVGLVGGACSDVAQTVLEM
                     RENGKMEKLLEILKSSEKERTMIFVNTKKKADFIAGYLCQEKFSSTSIHGDREQYQRE
                     SALWDFRTGKCTVIVCTAVAARGLDIENVQHVINYDVPKEVDEYVHRIGRTGRCGNTG
                     KATSFFNVQDDHVIARPLVKILTDAHQEVPAWLEEIAFGGHGALNSFYAADSMGEQAG
                     GNAVTTPSFAQEEEASWD"
     3'UTR           2169..2502
                     /gene="VLG1"
BASE COUNT      839 a    397 c    595 g    671 t
ORIGIN

  Query Match             32.8%;  Score 712;  DB 5;  Length 2502;
  Best Local Similarity   68.4%;  Pred. No. 1.6e-157;
  Matches 1001;  Conservative    0;  Mismatches  460;  Indels    3;  Gaps    1;

Qy        567 AAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGG 626
Db        593 AGGTGGTTTTAACAACTCTGATACTGGTGGAAGAGGACGAAGAGGTGGCCGAGGTGGTGG 652

Qy        627 TGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTC 686
Db        653 AAGCCAGTATGGTGGATACAAAGGAAGAAATGAAGAAGTTGGTGTAGAATCTGGAAAAAG 712

Qy        687 AGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCC 746
Db        713 TCAAGAAGAAGGAAATGAAAAAGATGAAA---AACCAAAAAAGGTGACCTATATTCCTCC 769

Qy        747 TCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGA 806
Db        770 GCCTCCACCTGATGGTGAAGATAATATATTCCGGCAATACCAGTCTGGAATCAATTTTGA 829

Qy        807 CAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGAC 866
Db        830 TAAATATGATGAGATTCTTGTTGATGTGACAGGAAAGGATGTTCCTCCTGCCATACTGAC 889

Qy        867 TTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATAC 926
Db        890 TTTTGAAGAAGCTAACCTTTGTGAAACACTAAGAAGAAATGTTGCTAGAGCTGGATATGT 949

Qy        927 TAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGAT 986
Db        950 AAAGCTAACACCAGTGCAGAAACACAGCATCCCTATTATAATGGCTGGTCGTGATTTAAT 1009

Qy        987 GGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCA 1046
Db       1010 GGCTTGCGCACAGACTGGTTCTGGTAAAACTGCTGCTTTTCTTTTGCCAATTCTCAGTTA 1069

Qy       1047 TATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTAT 1106
Db       1070 TATGATGAATGAAGGAATTACAGCTAGTCAGTATTTACAACTTCAGGAACCAGAAGCCAT 1129

Qy       1107 TATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTC 1166
Db       1130 AATTATTGCCCCTACCAGAGAACTTATTAACCAGATATATCTCGATGCTCGAAAATTTTC 1189

Qy       1167 TTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGGGACATTCAAT 1226
Db       1190 ATATGGAACCTGCGTGCGTCCAGTAGTTGTATATGGTGGTATACAACCTGTACATGCAAT 1249

Qy       1227 TCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATAT 1286
Db       1250 GAGGGACGTTGAAAAAGGGTGCAATATACTTTGTGCAACCCCTGGAAGATTGCTGGACAT 1309

Qy       1287 CATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGA 1346
Db       1310 AGTAAGCAAAGAAAAAATTGGCTTAAGTAAGCTAAGATATCTAGTTCTGGATGAAGCGGA 1369

Qy       1347 TCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAAT 1406
Db       1370 TCGCATGTTGGATATGGGATTTGCCCCGGAAATAGAAAAATTAATGACAAAGCCAGGAAT 1429

Qy       1407 GCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCA 1466
Db       1430 GCCAACAAAAGAAAAGCGACAAACACTAATGTTCAGTGCTACCTATCCTGAGGAAATTCG 1489

Qy       1467 AAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGG 1526
Db       1490 GAGGTTGGCTTCGAATTATTTGAAATCTGAACATTTGTTTGTTGTGGTCGGATTAGTTGG 1549

Qy       1527 TGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGA 1586
Db       1550 AGGAGCTTGTAGTGATGTGGCACAAACAGTTCTTGAAATGCGAGAAAATGGAAAGATGGA 1609

Qy       1587 AAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAAC 1646
Db       1610 AAAGCTACTTGAAATTCTGAAAAGCTCAGAGAAAGAGCGAACTATGATTTTTGTGAATAC 1669

Qy       1647 TAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAG 1706
Db       1670 AAAAAAGAAGGCAGATTTTATTGCTGGTTACCTTTGTCAAGAGAAATTTTCATCAACAAG 1729

Qy       1707 TATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGG 1766
Db       1730 CATTCACGGTGATAGAGAACAATACCAAAGAGAGAGTGCTCTCTGGGATTTCAGAACTGG 1789
```

```
Qy       1767 AAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGT 1826
Db       1790 AAAGTGTACTGTTATTGTCTGCACAGCAGTTGCTGCCAGAGGCTTGGATATTGAAAATGT 1849

Qy       1827 GCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGG 1886
Db       1850 TCAACACGTGATAAATTATGACGTTCCTAAGGAAGTTGATGAGTACGTCCATAGAATTGG 1909

Qy       1887 GCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGA 1946
Db       1910 TCGTACCGGTCGCTGTGGTAACACCGGAAAGGCAACATCATTTTTCAATGTTCAGGATGA 1969

Qy       1947 TAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGC 2006
Db       1970 CCATGTGATTGCTCGTCCCCTTGTGAAAATTCTTACCGATGCTCATCAAGAAGTCCCTGC 2029

Qy       2007 ATGGTTGGAAGAAATTGCCTTTAG 2030
Db       2030 TTGGTTAGAAGAAATTGCCTTTGG 2053
```

```
RESULT 9
AB005147
LOCUS       AB005147                2865 bp    mRNA    linear   VRT 01-OCT-1997
DEFINITION  Danio rerio mRNA for vas, complete cds.
ACCESSION   AB005147
VERSION     AB005147.1  GI:2463518
KEYWORDS    vas.
SOURCE      Danio rerio cDNA to mRNA.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (sites)
  AUTHORS   Yoon,C., Kawakami,K. and Hopkins,N.
  TITLE     Zebrafish vasa homologue RNA is localized to the cleavage planes of
            2- and 4-cell-stage embryos and is expressed in the primordial germ
            cells
  JOURNAL   Development 124 (16), 3157-3165 (1997)
  MEDLINE   97417583
REFERENCE   2  (bases 1 to 2865)
  AUTHORS   Kawakami,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-1997) Koichi Kawakami, The Institute of Medical
            Science, The University of Tokyo, Department of Tumor Biology;
            Shiroganedai 4-6-1, Minato-ku, Tokyo 108, Japan
            (E-mail:kawakami@ims.u-tokyo.ac.jp, Tel:81-3-5449-5308)
FEATURES             Location/Qualifiers
     source          1..2865
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
     CDS             71..2221
                     /codon_start=1
                     /product="vas"
                     /protein_id="BAA22535.1"
                     /db_xref="GI:2463519"
                     /translation="MDDWEEDQSPVVSCSSGFGLGSNGSDGGFKSFYTGGAGNDKSNS
                     EGTEGSSWKMTGDSFRGRGGRGGSRGGRGGFSGFKSEIDENGSDGGWNGGESRGRGRG
                     GFRGGFRSGSRDENDENRNDDGWKGGESRGRGRGGFGGSFRGGFRDGGNEDTGRRGFG
                     RENNENGNDEGGEGRGRGRGGFRGGFRDGGGDESGKRGFGRGGFRGRNEEVFSKVTTA
                     DKLDQEGSENAGPKVVYVPPPPPEEESSIFSHYATGINFDKYDDILVDVSGSNPPKAI
                     MTFEEAGLCDSLSKNVSKSGYVKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLP
                     ILQRFMTDGVAASKFSEIQEPEAIIVAPTRELINQIYLEARKFAYGTCVRPVVVYGGI
                     NTGYTIREVLKGCNVLCATPGRLHDLIGRGKIGLSKVRYLVLDEADRMLDMGFEPEMR
                     KLVASPGMPSKEERQTLMFSATYPEDIQRMAADFLKVDYIFLAVGVVGGACSDVEQTV
                     VQVDQYSKRDQLLELLRATGNERTMVFVETKRSADFIATFLCQEKISTTSIHGDREQR
                     EREKALSDFRLGHCPVLVATSVAARGLDIEQVQHVVNFDMPSSIDEYVHRIGRTGRCG
                     NTGRAVSFFNPESDTPLARSLVKVLSGAQQVVPKWLEEVAFSAHGTTGFNPRGKVFAS
                     TDSRKGGSFKSDEPPPSQTSAPSAAAAADDEEWE"
BASE COUNT      783 a    525 c    791 g    766 t
ORIGIN

  Query Match             31.6%;  Score 686.8;  DB 5;  Length 2865;
  Best Local Similarity   68.6%;  Pred. No. 1.4e-151;
  Matches  959;  Conservative    0;  Mismatches  436;  Indels    3;  Gaps    1;
```

```
Qy        702 AGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGA 761
Db        739 AGAAGGAAGTGAAAATGCAGGACCCAAGGTTGTTTATGTGCCGCCGCCTCCTCCAGAGGA 798

Qy        762 TGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTAT 821
Db        799 GGAGAGCTCGATATTCTCCCATTATGCAACAGGCATTAATTTTGACAAATATGATGATAT 858

Qy        822 TCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTGAAGAAGCTAA 881
Db        859 TCTTGTGGACGTGAGTGGCAGCAATCCTCCAAAAGCAATTATGACTTTTGAGGAAGCAGG 918

Qy        882 TCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGT 941
Db        919 ACTTTGTGACTCACTGAGCAAAAATGTATCAAAGTCTGGATATGTGAAGCCTACACCTGT 978

Qy        942 GCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAAC 1001
Db        979 TCAGAAACATGGCATTCCCATCATTTCAGCTGGACGGGATCTAATGGCTTGTGCTCAGAC 1038

Qy       1002 AGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGG 1061
Db       1039 TGGATCAGGAAAAACGGCGGCCTTCCTGCTGCCTATCCTACAGCGCTTTATGACTGATGG 1098

Qy       1062 AATAACTGCCAGTCGTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAAC 1121
Db       1099 TGTGGCAGCCAGCAAGTTCAGTGAGATACAGGAGCCTGAGGCCATAATCGTGGCTCCCAC 1158

Qy       1122 TCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGT 1181
Db       1159 CAGAGAACTTATCAATCAGATCTATCTGGAGGCCAGGAAGTTTGCATACGGGACCTGTGT 1218

Qy       1182 AAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACA 1241
Db       1219 ACGACCTGTTGTGGTTTATGGAGGTATAAATACTGGATATACTATCCGAGAGGTGTTAAA 1278

Qy       1242 AGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAA 1301
Db       1279 GGGCTGCAATGTTCTGTGTGCTACTCCTGGAAGATTGCATGACCTCATTGGTCGTGGAAA 1338

Qy       1302 GATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATGTTGGATAT 1361
Db       1339 GATTGGCCTGAGTAAAGTGCGCTATCTAGTTCTGGATGAAGCAGACAGAATGCTGGACAT 1398

Qy       1362 GGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACA 1421
Db       1399 GGGCTTTGAGCCAGAAATGCGCAAACTGGTGGCGTCTCCTGGTATGCCTTCAAAAGAGGA 1458

Qy       1422 GCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGA 1481
Db       1459 ACGACAAACCCTCATGTTCAGTGCCACCTACCCTGAAGATATTCAAAGAATGGCTGCTGA 1518

Qy       1482 GTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGA 1541
Db       1519 CTTTCTTAAAGTGGACTATATTTTCCTTGCTGTTGGTGTGGTGGGTGGAGCATGCAGTGA 1578

Qy       1542 TGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAAT 1601
Db       1579 TGTGGAGCAAACCGTTGTTCAGGTGGACCAGTACTCAAAGAGGGACCAGCTGCTTGAATT 1638

Qy       1602 TCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGA 1661
Db       1639 GCTCAGAGCAACAGGTAATGAGCGCACAATGGTTTTTGTGGAAACCAAAAGAAGTGCTGA 1698

Qy       1662 TTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCG 1721
Db       1699 TTTCATAGCAACATTTCTCTGTCAAGAGAAGATCTCAACCACAAGCATCCATGGTGATCG 1758

Qy       1722 GGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCT 1781
```

```
Db       1759 GGAACAGCGAGAGCGAGAGAAAGCTCTCAGTGATTTTCGCCTTGGCCACTGTCCGGTTCT 1818

Qy       1782 TGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAA 1841

Db       1819 GGTCGCCACATCTGTAGCTGCTAGAGGCCTGGATATTGAGCAAGTCCAGCATGTTGTGAA 1878

Qy       1842 TTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTG 1901

Db       1879 TTTTGACATGCCCAGCAGCATCGATGAGTATGTCCATCGCATAGGAAGAACTGGACGCTG 1938

Qy       1902 TGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACA 1961

Db       1939 TGGGAACACCGGTCGCGCCGTGTCCTTTTTTAACCCAGAGTCTGACACTCCACTAGCTCG 1998

Qy       1962 GCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAAT 2021

Db       1999 CTCCCTGGTCAAAGTCCTTTCAGGGGCCCAACAAGTAGTTCCAAAATGGCTGGAGGAAGT 2058

Qy       2022 TGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATC 2081

Db       2059 TGCCTTCAGTGCTCA---TGGCACAACAGGCTTCAACCCACGTGGAAAGGTGTTTGCATC 2115

Qy       2082 AGTTGATACCAGAAAGGG 2099

Db       2116 TACAGACTCGCGGAAGGG 2133


RESULT 10
DRY12007
LOCUS       DRY12007               2492 bp    mRNA    linear   VRT 09-OCT-1997
DEFINITION  D.rerio vlg (vasa like gene) mRNA for putative RNA helicase.
ACCESSION   Y12007
VERSION     Y12007.1  GI:2558534
KEYWORDS    RNA helicase; vasa gene; vlg gene.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 2492)
  AUTHORS   Olsen,L.C., Aasland,R. and Fjose,A.
  TITLE     A vasa-like gene in zebrafish identifies putative primordial germ
            cells
  JOURNAL   Mech. Dev. 66 (1-2), 95-105 (1997)
  MEDLINE   98025484
   PUBMED   9376327
REFERENCE   2  (bases 1 to 2492)
  AUTHORS   Olsen,L.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-1997) L.C. Olsen, University of Bergen,
            Department of Molecular Biology, Thormoehlensgt 55, N-5020 Bergen,
            NORWAY
FEATURES             Location/Qualifiers
     source          1..2492
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /dev_stage="embryonic, late blastula"
     gene            36..2138
                     /gene="vlg"
     CDS             36..2138
                     /gene="vlg"
                     /note="vasa-like gene; putative"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="RNA helicase (DEAD box)"
                     /protein_id="CAA72735.1"
                     /db_xref="GI:2558535"
                     /db_xref="SPTREMBL:O42378"
                     /translation="MDDWEEDQSPVVSCSSGFGGAGNDKSNSEGTEGSSWKMTGDSFR
                     GRGGRGGSRGGRGGFSGFKSEIDENGSDGGWNGGESRGRGRGGFRGGFRSGSRDENDE
                     NGNDDGWKGGESRGRGRGGFGGGFRGGFRDGGNEDTGRRGFGRENNENGNDEGGEGRG
                     RGRGGFRGGFRDGGGDESGKRGFGRGGFRGRNEEVFSKVTTADKLDQEGSENAGPKVV
                     YVPPPPPEEESSIFSHYATGINFDKYDDILVDVSGSNPPKAIMTFEEAGLCDSLSKNV
                     SKSGYVKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQRFMTDGVAASKFS
                     EIQEPEAIIVAPTRELINQIYLEARKFAYGTCVRPVVVYGGINTGYTIREVLKGCNVL
                     CATPGRLHDLIGRGKIGLSKVRYLVLDEADRMLDMGFEPEMRKLVASPGMPSKEKRQT
                     LMFSATYPEDIQRMAADFLKVDYIFLAVGVVGGACSDVEQTIVQVDQYSKRDQLLELL
                     RATGNERTMVFVETKRSADFIATFLCQEKISTTSIHGDREQREREKALSDFRLGHCPV
                     LVATSVAARGLDIEQVQHVVNFDMPSSIDEYVHRIGRTGRCGNTGRAVSFFNPESDTP
                     LARSLVKVLSGAQQVVPKWLEEVAFSAHGTTGFNPRGKVFASTDSRKGGSFKSDEPPP
                     SQTSAPSAAAAADDEEWE"
     protein_bind    2255..2277
                     /note="similar to nanos response element"
                     /bound_moiety="a putative nanos-dependent factor"
                     /evidence=not_experimental
BASE COUNT      703 a    453 c    703 g    633 t
ORIGIN

  Query Match             31.6%;  Score 686.4;  DB 5;  Length 2492;
  Best Local Similarity   69.6%;  Pred. No. 1.8e-151;
  Matches  927;  Conservative    0;  Mismatches  405;  Indels    0;  Gaps    0;

Qy        702 AGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGA 761

Db        656 AGAAGGAAGTGAAAATGCAGGACCCAAGGTTGTTTATGTGCCGCCGCCTCCTCCAGAGGA 715

Qy        762 TGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTAT 821

Db        716 GGAGAGCTCGATATTCTCCCATTATGCAACAGGCATTAATTTTGACAAATATGATGATAT 775

Qy        822 TCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAA 881

Db        776 TCTTGTGGACGTGAGTGGCAGCAATCCTCCAAAAGCAATTATGACTTTTGAGGAAGCAGG 835

Qy        882 TCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGT 941

Db        836 ACTTTGTGACTCACTGAGCAAAAATGTATCAAAGTCTGGATATGTGAAGCCTACACCTGT 895

Qy        942 GCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAAC 1001

Db        896 TCAGAAACATGGCATTCCCATCATTTCAGCTGGACGGGATCTAATGGCTTGTGCTCAGAC 955

Qy       1002 AGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGG 1061

Db        956 TGGATCAGGAAAAACGGCGGCCTTCCTGCTGCCCATCCTACAGCGCTTTATGACTGATGG 1015

Qy       1062 AATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAAC 1121

Db       1016 TGTGGCAGCCAGCAAGTTCAGTGAGATACAGGAGCCTGAGGCCATAATCGTGGCTCCCAC 1075

Qy       1122 TCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGT 1181

Db       1076 CAGAGAACTTATCAATCAGATCTATCTGGAGGCCAGGAAGTTTGCATACGGGACCTGTGT 1135

Qy       1182 AAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACA 1241

Db       1136 ACGACCTGTTGTGGTTTATGGAGGTATAAATACCGGATATACTATCCGAGAGGTGTTAAA 1195

Qy       1242 AGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAA 1301

Db       1196 GGGCTGCAATGTATTGTGTGCTACTCCTGGAAGATTGCATGACCTCATTGGTCGTGGAAA 1255

Qy       1302 GATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATAT 1361

Db       1256 GATTGGCCTGAGTAAAGTGCGCTATCTAGTTCTGGATGAAGCAGACAGAATGCTGGACAT 1315

Qy       1362 GGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACA 1421

Db       1316 GGGCTTTGAGCCAGAAATGCGCAAACTGGTGGCGTCTCCTGGTATGCCTTCAAAAGAGAA 1375

Qy       1422 GCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGA 1481

Db       1376 ACGACAAACCCTCATGTTCAGTGCCACCTACCCTGAAGATATTCAAAGAATGGCTGCTGA 1435

Qy       1482 GTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGA 1541
```

```
Db     1436 CTTTCTTAAAGTGGACTATATTTTCCTTGCTGTTGGTGTGGTGGGTGGAGCATGCAGTGA 1495

Qy     1542 TGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAAT 1601
Db     1496 TGTGGAACAAACCATTGTTCAGGTGGACCAGTACTCAAAGAGGGACCAGCTGCTTGAATT 1555

Qy     1602 TCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGA 1661
Db     1556 GCTCAGAGCAACAGGTAATGAGCGCACAATGGTTTTTGTTGAAACCAAAAGAAGTGCTGA 1615

Qy     1662 TTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCG 1721
Db     1616 TTTCATAGCAACATTTCTCTGTCAAGAGAAGATCTCAACCACAAGCATCCATGGTGATCG 1675

Qy     1722 GGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTCGCTTTGGAAAGTGCCCAGTTCT 1781
Db     1676 GGAACAGCGAGAGCGAGAGAAAGCTCTCAGTGATTTCGCCTTGGCCACTGTCCGGTTTT 1735

Qy     1782 TGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAA 1841
Db     1736 GGTCGCCACATCTGTAGCTGCTAGAGGCCTGGATATTGAGCAAGTCCAGCATGTTGTGAA 1795

Qy     1842 TTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTG 1901
Db     1796 TTTTGACATGCCCAGCAGCATCGATGAGTATGTCCATCGCATAGGAAGAACTGGACGCTG 1855

Qy     1902 TGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACA 1961
Db     1856 TGGGAACACCGGTCGGGCCGTGTCCTTTTTTAACCCAGAGTCTGACACTCCACTAGCTCG 1915

Qy     1962 GCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAAT 2021
Db     1916 CTCCCTGGTCAAAGTCCTTTCAGGGGCCCAACAAGTAGTTCCAAAATGGCTGGAGGAAGT 1975

Qy     2022 TGCCTTTAGTAC 2033
Db     1976 TGCCTTCAGTGC 1987


RESULT 11
AB051835
LOCUS       AB051835                2058 bp    mRNA    linear   VRT 14-FEB-2002
DEFINITION  Oreochromis niloticus vas-s mRNA for vasa short form, complete cds.
ACCESSION   AB051835
VERSION     AB051835.1  GI:14211583
KEYWORDS    .
SOURCE      Oreochromis niloticus cDNA to mRNA, clone_lib:lamda zapII tilapia
            ovary cDNA clone:vas-s.
  ORGANISM  Oreochromis niloticus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
            Labroidei; Cichlidae; Oreochromis.
REFERENCE   1
  AUTHORS   Kobayashi,T., Kajiura-Kobayashi,H. and Nagahama,Y.
  TITLE     Two isoforms of vasa homologs in a teleost fish: their differential
            expression during germ cell differentiation
  JOURNAL   Mech. Dev. 111 (1-2), 167-171 (2002)
  MEDLINE   21664110
REFERENCE   2  (bases 1 to 2058)
  AUTHORS   Kobayashi,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-NOV-2000) Tohru Kobayashi, National Institute for
            Basic Biology, Department of Developmental Biology; Myodaiji-cho,
            Okazaki, Aichi 444-8585, Japan (E-mail:tohru_k@nibb.ac.jp,
            Tel:81-564-55-7554, Fax:81-564-55-7556)
FEATURES             Location/Qualifiers
     source          1..2058
                     /organism="Oreochromis niloticus"
                     /db_xref="taxon:8128"
                     /clone="vas-s"
                     /clone_lib="lamda zapII tilapia ovary cDNA"
     gene            1..2058
                     /gene="vas-s"
     CDS             48..1913
                     /gene="vas-s"
                     /codon_start=1
                     /product="vasa short form"
                     /protein_id="BAB56110.1"
                     /db_xref="GI:14211584"
                     /translation="MDEWEEQETATSGTCALTSDTSTGGTQGDFWNSDHGEFQGGRGG
                     RGRGRGGFKNSFFSDGDGQNNTAERSSFSRRGGRANEEDGDNKNVSRGRRRGRGGGEQ
                     GGRSGFGGGYRGKDEKSLLKARGYPAENADASDRPRVTYVPPPPSEDEESIFSHYESG
                     VNFDKYDEILVDVSGTNPPPAIMTFDEAALCESLKRNVSKSGYVKPTPVQKHGIPIIS
                     AGRDLMACAQTGSGKTAAFLLPILQQLMADGAAASCFSEMQEPDAIIVAPTRELINQI
                     YLEARKFAYGTCVRPVVVYGGVSTGHQIRDLLRGCNVLCGTPGRLLDMIGRGKVGLTK
                     VRYLVLDEADRMLDMGFEPDMRRLVGSPGMPSKENRQTLMFSATFPEDIQRLAADFLK
                     TDYLFLAVGIVGGACSDVEQTFVQVTKFAKREQLIDLLKSTGSERTMVFVETKRQADF
                     IATILCQEKFPTTSIHGDREQWQREQALGDFRSGKCSVLVATSVGARGLDIPDVQFVV
                     NFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPGADSELARSLVTILSKAQQEVPSWL
                     EEFAFSGHSPSGFNPPRKHFASTDLRKGPQGGFFQDGSVTSQLAAQTAADDEWE"
BASE COUNT      542 a    447 c    596 g    473 t
ORIGIN

  Query Match             28.5%;  Score 619.4;  DB 5;  Length 2058;
  Best Local Similarity   66.8%;  Pred. No. 1.2e-135;
  Matches  882;  Conservative    0;  Mismatches  438;  Indels    0;  Gaps    0;

Qy      711 TGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTC 770
Db      443 TGCGAGTGATAGACCCAGGGTTACCTACGTCCCCCCACCACCCTCTGAAGATGAGGAGTC 502

Qy      771 CATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGA 830
Db      503 CATTTTTTCCCACTATGAGTCCGGTGTCAACTTTGACAAGTACGATGAGATCTTGGTGGA 562

Qy      831 AGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCA 890
Db      563 TGTCAGTGGAACCAACCCACCACCAGCTATAATGACATTTGATGAAGCAGCGCTGTGCGA 622

Qy      891 GACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATA 950
Db      623 GTCCCTAAAGAGAAACGTCAGCAAGTCTGGTTATGTGAAGCCCACCCCTGTGCAGAAGCA 682

Qy      951 CAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGG 1010
Db      683 CGGCATCCCCATCATTTCTGCTGGCAGGGATCTCATGGCCTGCGCCCAGACTGGGTCTGG 742

Qy     1011 GAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGC 1070
Db      743 AAAAACGGCTGCCTTCCTGCTCCCAATCCTTCAACAGCTGATGGCAGATGGTGCGGCGGC 802

Qy     1071 CAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATT 1130
Db      803 CAGTTGCTTCAGTGAGATGCAGGAGCCTGATGCCATCATTGTGGCACCAACTAGGGAGCT 862

Qy     1131 GGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGT 1190
Db      863 GATCAACCAGATTTACCTGGAGGCCAGGAAGTTTGCCTATGGGACATGTGTGCGACCTGT 922

Qy     1191 TGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAA 1250
Db      923 TGTAGTGTATGGTGGAGTCAGCACTGGACACCAAATAAGAGACCTCTTGAGGGGCTGTAA 982

Qy     1251 TATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCT 1310
Db      983 TGTGTTGTGTGGAACACCAGGACGACTGCTGGACATGATTGGAAGAGGAAAGGTTGGCTT 1042

Qy     1311 CAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGG 1370
Db     1043 GACTAAAGTGCGTTACTTGGTGCTGGATGAGGCTGACCGCATGTTGGATATGGGGTTTGA 1102

Qy     1371 TCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAAC 1430
Db     1103 GCCTGACATGCGCCGCCTGGTGGGCTCTCCTGGGATGCCATCAAAAGAGAACCGTCAGAC 1162
```

```
Qy       1431 CCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAA 1490
Db       1163 TCTGATGTTCAGTGCTACTTTCCCCGAGGACATCCAGAGGCTTGCAGCTGACTTTCTAAA 1222

Qy       1491 GTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCA 1550
Db       1223 GACTGACTATTTGTTCTTGGCCGTGGGTATTGTGGGTGGAGCCTGCAGTGATGTGGAGCA 1282

Qy       1551 GACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAA 1610
Db       1283 GACATTTGTGCAGGTCACCAAGTTCGCAAAGAGGGAGCAGCTAATAGACCTCCTGAAATC 1342

Qy       1611 CATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGC 1670
Db       1343 AACAGGATCGGAGCGCACCATGGTGTTTGTAGAGACCAAGAGACAGGCTGATTTTATTGC 1402

Qy       1671 AACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAG 1730
Db       1403 CACAATCTTGTGTCAGGAGAAGTTTCCAACTACCAGCATTCACGGTGACCGTGAGCAGTG 1462

Qy       1731 AGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTAC 1790
Db       1463 GCAGCGAGAGCAGGCCCTTGGAGACTTCCGCTCTGGCAAATGTTCTGTCCTGGTGGCAAC 1522

Qy       1791 TTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCT 1850
Db       1523 CTCTGTAGGTGCCCGTGGACTCGATATTCCAGATGTGCAGTTCGTGGTGAACTTTGACCT 1582

Qy       1851 TCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATAC 1910
Db       1583 CCCAAACAATATCGACGAATATGTCCATCGTATTGGTAGAACCGGCCGCTGTGGCAACAC 1642

Qy       1911 TGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGT 1970
Db       1643 CGGGAGGGCAGTGTCCTTCTATGATCCAGGAGCTGACAGTGAGTTGGCTCGATCACTGGT 1702

Qy       1971 AAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAG 2030
Db       1703 CACAATCCTGTCCAAGGCCCAGCAGGAGGTTCCCTCGTGGTTAGAGGAGTTTGCGTTCAG 1762

RESULT 12
AB032467
LOCUS       AB032467                2130 bp    mRNA    linear   VRT 27-DEC-2000
DEFINITION  Oreochromis niloticus mRNA for vasa, complete cds.
ACCESSION   AB032467
VERSION     AB032467.1  GI:11990881
KEYWORDS    vasa; vas.
SOURCE      Oreochromis niloticus cDNA to mRNA.
  ORGANISM  Oreochromis niloticus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
            Labroidei; Cichlidae; Oreochromis.
REFERENCE   1  (sites)
  AUTHORS   Kobayashi,T., Kajiura-Kobayashi,H. and Nagahama,Y.
  TITLE     Differential expression of vasa homologue gene in the germ cells
            during oogenesis and spermatogenesis in a teleost fish, tilapia,
            Oreochromis niloticus
  JOURNAL   Mech. Dev. 99 (1-2), 139-142 (2000)
  MEDLINE   20544961
REFERENCE   2  (bases 1 to 2130)
  AUTHORS   Kobayashi,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-1999) Tohru Kobayashi, National Institute for
            Basic Bilogy, Laboratory of Reproductive Biology; 38 Nishiginaka,
            Myodaiji-cho, Okazaki, Aichi 444-8585, Japan
            (E-mail:tohru_k@nibb.ac.jp, Tel:81-564-55-7554, Fax:81-564-55-7556)
FEATURES             Location/Qualifiers
     source          1..2130
                     /organism="Oreochromis niloticus"
                     /db_xref="taxon:8128"
     gene            1..2130
                     /gene="vas"
     CDS             48..1985
                     /gene="vas"
                     /codon_start=1
                     /product="vasa"
                     /protein_id="BAB19807.1"
                     /db_xref="GI:11990882"
                     /translation="MDEWEEQETATSGTCALTSDTSTGGTQGDFWNSDHGEFQGGRGG
                     RGRGGGFKNSFFSGAEDGANDGDGDGQNNTAERSSFSRRGGRGRGKGFGRTDHSDFD
                     ANEEDGDNKNVSRGRRRGRGGGEQGGRSGFGGGYRGKDEKSLLKARGYPAENADASDR
                     PRVTYVPPPPSEDEESIFSHYESGVNFDKYDEILVDVSGTNPPPAIMTFDEAALCESL
                     KRNVSKSGYVKPTPVQKHGIPIISAGRDLMACAQTGSGKTAAFLLPILQQLMADGAAA
                     SCFSEMQEPDAIIVAPTRELINQIYLEARKFAYGTCVRPVVVYGGVSTGHQIRDLLRG
                     CNVLCGTPGRLLDMIGRGKVGLTKVRYLVLDEADRMLDMGFEPDMRRLVGSPGMPSKE
                     NRQTLMFSATFPEDIQRLAADFLKTDYLFLAVGIVGGACSDVEQTFVQVTKFAKREQL
                     IDLLKSTGSERTMVFVETKRQADFIATILCQEKFPTTSIHGDREQWQREQALGDFRSG
                     KCSVLVATSVGARGLDIPDVQFVVNFDLPNNIDEYVHRIGRTGRCGNTGRAVSFYDPG
                     ADSELARSLVTILSKAQQEVPSWLEEFAFSGHSPSGFNPPRKHFASTDLRKGPQGGFF
                     QDGSVTSQLAAQTAADDEEWE"
BASE COUNT      560 a    458 c    624 g    488 t
ORIGIN

  Query Match             28.5%;  Score 619.4;  DB 5;  Length 2130;
  Best Local Similarity   66.8%;  Pred. No. 1.2e-135;
  Matches  882;  Conservative    0;  Mismatches  438;  Indels    0;  Gaps    0;

Qy        711 TGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTC 770
Db        515 TGCGAGTGATAGACCCAGGGTTACCTACGTCCCCCCACCACCCTCTGAAGATGAGGAGTC 574

Qy        771 CATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGA 830
Db        575 CATTTTTTCCCACTATGAGTCCGGTGTCAACTTTGACAAGTACGATGAGATCTTGGTGGA 634

Qy        831 AGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCA 890
Db        635 TGTCAGTGGAACCAACCCACCACCAGCTATAATGACATTTGATGAAGCAGCGCTGTGCGA 694

Qy        891 GACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATA 950
Db        695 GTCCCTAAAGAGAAACGTCAGCAAGTCTGGTTATGTGAAGCCCACCCCTGTGCAGAAGCA 754

Qy        951 CAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGG 1010
Db        755 CGGCATCCCCATCATTTCTGCTGGCAGGGATCTCATGGCCTGCGCCCAGACTGGGTCTGG 814

Qy       1011 GAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGC 1070
Db        815 AAAAACGGCTGCCTTCCTGCTCCCAATCCTTCAACAGCTGATGGCAGATGGTGCGGCGGC 874

Qy       1071 CAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATT 1130
Db        875 CAGTTGCTTCAGTGAGATGCAGGAGCCTGATGCCATCATTGTGGCACCAACTAGGGAGCT 934

Qy       1131 GGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGT 1190
Db        935 GATCAACCAGATTTACCTGGAGGCCAGGAAGTTTGCCTATGGGACATGTGTGCGACCTGT 994

Qy       1191 TGTTATATATGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAA 1250
Db        995 TGTAGTGTATGGTGGAGTCAGCACTGGACACCAAATAAGAGACCTCTTGAGGGGCTGTAA 1054

Qy       1251 TATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCT 1310
Db       1055 TGTGTTGTGTGGAACACCAGGACGACTGCTGGACATGATTGGAAGAGGAAAGGTTGGCTT 1114

Qy       1311 CAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGG 1370
Db       1115 GACTAAAGTGCGTTACTTGGTGCTGGATGAGGCTGACCGCATGTTGGATATGGGGTTTGA 1174

Qy       1371 TCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAAC 1430
Db       1175 GCCTGACATGCGCCGCCTGGTGGGCTCTCCTGGGATGCCATCAAAAGAGAACCGTCAGAC 1234
```

```
Qy       1431 CCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAA 1490
Db       1235 TCTGATGTTCAGTGCTACTTTCCCCGAGGACATCCAGAGGCTTGCAGCTGACTTTCTAAA 1294

Qy       1491 GTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCA 1550
Db       1295 GACTGACTATTTGTTCTTGGCCGTGGGTATTGTGGGTGGAGCCTGCAGTGATGTGGAGCA 1354

Qy       1551 GACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAA 1610
Db       1355 GACATTTGTGCAGGTCACCAAGTTCGCAAAGAGGGAGCAGCTAATAGACCTCCTGAAATC 1414

Qy       1611 CATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGC 1670
Db       1415 AACAGGATCGGAGCGCACCATGGTGTTTGTAGAGACCAAGAGACAGGCTGATTTTATTGC 1474

Qy       1671 AACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAG 1730
Db       1475 CACAATCTTGTGTCAGGAGAAGTTTCCAACTACCAGCATTCACGGTGACCGTGAGCAGTG 1534

Qy       1731 AGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTAC 1790
Db       1535 GCAGCGAGAGCAGGCCCTTGGAGACTTCCGCTCTGGCAAATGTTCTGTCCTGGTGGCAAC 1594

Qy       1791 TTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCT 1850
Db       1595 CTCTGTAGGTGCCCGTGGACTCGATATTCCAGATGTGCAGTTCGTGGTGAACTTTGACCT 1654

Qy       1851 TCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATAC 1910
Db       1655 CCCAAACAATATCGACGAATATGTCCATCGTATTGGTAGAACCGGCCGCTGTGGCAACAC 1714

Qy       1911 TGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGT 1970
Db       1715 CGGGAGGGCAGTGTCCTTCTATGATCCAGGAGCTGACAGTGAGTTGGCTCGATCACTGGT 1774

Qy       1971 AAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAG 2030
Db       1775 CACAATCCTGTCCAAGGCCCAGCAGGAGGTTCCCTCGTGGTTAGAGGAGTTTGCGTTCAG 1834
```

RESULT 13
AB032566

```
LOCUS       AB032566                2612 bp    mRNA    linear   VRT 06-JUL-2001
DEFINITION  Oncorhynchus mykiss vas mRNA for Vasa, complete cds.
ACCESSION   AB032566
VERSION     AB032566.1  GI:6521013
KEYWORDS    Vasa.
SOURCE      Oncorhynchus mykiss female ovary cDNA to mRNA.
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1
  AUTHORS   Yoshizaki,G., Sakatani,S., Tominaga,H. and Takeuchi,T.
  TITLE     Cloning and characterization of a vasa-like gene in rainbow trout
            and its expression in the germ cell lineage
  JOURNAL   Mol. Reprod. Dev. 55 (4), 364-371 (2000)
  MEDLINE   20160526
REFERENCE   2  (bases 1 to 2612)
  AUTHORS   Yoshizaki,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-SEP-1999) Goro Yoshizaki, Tokyo University of
            Fisheries, Department of Aquatic Biosciences; Konan, 4-5-7, Minato,
            Tokyo 108-8477, Japan (E-mail:goro@tokyo-u-fish.ac.jp,
            Tel:81-3-5463-0558, Fax:81-3-5463-0558)
FEATURES             Location/Qualifiers
     source          1..2612
                     /organism="Oncorhynchus mykiss"
                     /db_xref="taxon:8022"
                     /sex="female"
                     /tissue_type="ovary"
     gene            1..2612
                     /gene="vas"
     CDS             56..1999
                     /gene="vas"
                     /codon_start=1
                     /product="Vasa"
                     /protein_id="BAA88059.1"
                     /db_xref="GI:6521014"
                     /translation="MDNWDEVTTSNESGLSSWTSGGQSSSFGRPTDKVSSWNSGCGGF
                     GGGRGRGSRGGGGDFKSFSSGAGENGNEDKDSSWNSGGGFRGRGRGGGRGSRGGIRN
                     GDDGYDGGFGGSQGGRGGRGGFRSGGGDGGGFGGGGYRGRDEEIFSKGSTMDGGGDGG
                     NPGPPKVTYVPEALCEKESSIFAHYESGINFDKYDDILVDVSGSNPPKAIMGFEEAAL
                     CESLNRNISKSGYKKPTPVQKHGIPIIAAGRDLMACAQTGSGKTAAFLLPILQQLMVD
                     GVAASQFSEIQEPEVIIVAPTRELINQIYMEARKFAHGTCVRPVVVYGGISTGHTIRE
                     ILKGCNVLCATPGRLMDIIGRGKIGLSKLRYLVLDEADRMLDMGFEPAMRKLVGSPGI
                     PAKEDRQTLMFSATYPEDIQKLAADFLKKDYLFLAVGVVGGACSDVEQVVVQVTKFSK
                     RDQLLEVLKTTGSERTMVFVETKRQADFIATFLCREKVNTTSIHGDREQREREQALGD
                     FRSGRCPVLVATSVAARGLDIKDVQHIVNFDLPNNIDEYVHRIGRTGRCGNTGRAVCF
                     FDPGADGNLARSLVKVLSGAQQEVPKWLEEAAFSAHGTTGFNPGGRTFASTDSRKGGS
                     LQRDGASHLAAALSSGADDDEWE"
BASE COUNT      625 a    548 c    775 g    664 t
ORIGIN

  Query Match             28.2%;  Score 612.2;  DB 5;  Length 2612;
  Best Local Similarity   66.1%;  Pred. No. 6.1e-134;
  Matches  884;  Conservative    0;  Mismatches  453;  Indels    0;  Gaps    0;

Qy        697 GGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCT 756
Db        521 GGTGGAGATGGTGGCAATCCAGGTCCCCCAAAGGTGACCTATGTGCCCGAAGCCCTCTGT 580

Qy        757 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGAC 816
Db        581 GAGAAGGAGTCGTCTATTTTTGCTCATTATGAGTCAGGCATCAACTTTGATAAGTACGAT 640

Qy        817 ACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAA 876
Db        641 GACATCTTGGTCGATGTCAGCGGCAGCAATCCGCCAAAGGCAATCATGGGTTTTGAAGAG 700

Qy        877 GCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACT 936
Db        701 GCTGCCCTGTGTGAATCTCTGAACAGGAACATCAGTAAGTCTGGCTATAAGAAGCCCACC 760

Qy        937 CCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCT 996
Db        761 CCAGTACAGAAGCATGGCATTCCCATCATCGCTGCTGGCCGGGACCTCATGGCCTGTGCC 820

Qy        997 CAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCAT 1056
Db        821 CAGACTGGATCTGGGAAAACGGCTGCATTCTTACTGCCCATCCTGCAGCAGCTGATGGTG 880

Qy       1057 GATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCA 1116
Db        881 GACGGCGTGGCAGCTAGTCAGTTCAGCGAGATCCAGGAGCCAGAGGTTATTATAGTGGCC 940

Qy       1117 CCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACT 1176
Db        941 CCAACCAGGGAGCTGATCAACCAGATCTACATGGAAGCCCGGAAGTTTGCCCATGGGACT 1000

Qy       1177 TGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATA 1236
Db       1001 TGTGTGCGTCCTGTGGTGGTTTACGGTGGCATAAGCACCGGTCACACCATTCGCGAGATA 1060

Qy       1237 GTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAA 1296
Db       1061 CTGAAGGGCTGCAATGTGCTGTGTGCGACTCCAGGAAGACTGATGGACATTATTGGAAGA 1120

Qy       1297 GAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTG 1356
Db       1121 GGAAAGATTGGCTTGAGCAAGCTGCGTTACCTGGTGCTGGACGAGGCTGACAGGATGCTG 1180

Qy       1357 GATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAG 1416
```

```
Db      1181 GACATGGGCTTTGAGCCTGCGATGCGCAAGCTGGTGGGGTCCCCAGGCATACCAGCTAAA 1240

Qy      1417 GAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCT 1476
Db      1241 GAGGATCGCCAGACCCTTATGTTCAGCGCTACCTACCCCGAGGACATCCAAAAACTGGCT 1300

Qy      1477 GCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGT 1536
Db      1301 GCTGACTTCCTGAAGAAGGACTATCTGTTCCTTGCTGTGGGTGTGGTGGGCGGAGCCTGC 1360

Qy      1537 AGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTT 1596
Db      1361 AGTGATGTGGAGCAGGTCGTGGTTCAGGTGACCAAGTTCTCCAAAAGAGACCAACTACTG 1420

Qy      1597 GAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAA 1656
Db      1421 GAGGTCCTGAAGACTACAGGGTCTGAACGCACAATGGTCTTTGTGGAAACCAAGAGGCAG 1480

Qy      1657 GCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGT 1716
Db      1481 GCTGACTTTATAGCAACGTTCCTGTGTCGGGAGAAGGTTAATACTACTAGTATTCACGGT 1540

Qy      1717 GATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCA 1776
Db      1541 GACCGTGAGCAGAGGGAGCGTGAACAGGCGCTCGGCGACTTCCGCTCCGGAAGGTGTCCT 1600

Qy      1777 GTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTT 1836
Db      1601 GTCCTGGTGGCCACCTCTGTTGCTGCCCGCGGACTGGACATCAAGGATGTCCAACACATA 1660

Qy      1837 ATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGT 1896
Db      1661 GTCAACTTTGACCTCCCCAACAACATTGATGAGTACGTCCACCGCATCGGGAGAACAGGT 1720

Qy      1897 CGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTA 1956
Db      1721 CGCTGTGGGAACACCGGGAGAGCTGTGTGTTTCTTTGACCCGGGGGCTGACGGCAACCTG 1780

Qy      1957 GCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAA 2016
Db      1781 GCGCGCTCCCTGGTCAAAGTCTTGTCTGGGGCCCAGCAGGAGGTGCCCAAATGGCTGGAG 1840

Qy      2017 GAAATTGCCTTTAGTAC 2033
Db      1841 GAAGCTGCCTTCAGTGC 1857
```

```
RESULT 14
AF479823
LOCUS       AF479823                 1512 bp    mRNA    linear   VRT 01-APR-2002
DEFINITION  Pantodon buchholzi DEAD box RNA helicase Vasa mRNA, complete cds.
ACCESSION   AF479823
VERSION     AF479823.1  GI:19387222
KEYWORDS    .
SOURCE      Pantodon buchholzi.
  ORGANISM  Pantodon buchholzi
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
            Osteoglossiformes; Pantodontidae; Pantodon.
REFERENCE   1  (bases 1 to 1512)
  AUTHORS   Knaut,H., Steinbeisser,H., Schwarz,H. and Nusslein-Volhard,C.
  TITLE     An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA
            Translation to the Germ Cells in the Zebrafish
  JOURNAL   Curr. Biol. 12 (6), 454-466 (2002)
  MEDLINE   21906633
   PUBMED   11909530
REFERENCE   2  (bases 1 to 1512)
  AUTHORS   Knaut,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-2002) Genetik, Max-Planck Institut fuer
            Entwicklungsbiologie, Spemannstrasse 35, Tuebingen 72076, Germany
FEATURES             Location/Qualifiers
     source          1..1512
                     /organism="Pantodon buchholzi"
                     /db_xref="taxon:8276"
     CDS             1..1185
                     /codon_start=1
                     /product="DEAD box RNA helicase Vasa"
                     /protein_id="AAL87142.1"
                     /db_xref="GI:19387223"
                     /translation="MACAQTGSGKTAAFLLPILQQLMTSGVASSRFSEVQEPEAIVVA
                     PTRELINQIYLEARKFAFGTCVRPVVVYGGTSTMYTIREVLKGCNILCGTPGRLLDII
                     GRGKVGLNKLRYLVLDEADRMLDMGFEPDMRRLVASPGMPPKEERQTLMFSATYPEDI
                     QRLAADFLKRDYLFLVVGQVGGACNDVEQHVVEVSQYSKREQLVEILRTTGNERTMVF
                     VETKRKADFIATFLCQEQISTTSIHGDREQREREQALGDFRSGKCPVLVATSVAARGL
                     DIEHVQHVVNFDLPSTIDEYVHRIGRTGRCGNTGKAISFFDPEADAPLARSLVKILSG
                     AQQEVPKWLEEIAFSAHGTTAFNPRGKVFASTDTRKEGAFPRAEPVQPAVQNPDADEE
                     WE"
BASE COUNT      387 a    301 c    406 g    415 t      3 others
ORIGIN

  Query Match             26.0%;  Score 565;  DB 5;  Length 1512;
  Best Local Similarity   71.0%;  Pred. No. 8.9e-123;
  Matches  745;  Conservative    2;  Mismatches  302;  Indels    0;  Gaps    0;

Qy       985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCT 1044
Db         1 ATGGCATGYGCTCARACTGGATCGGGTAAGACGGCTGCATTCTTATTACCCATCCTGCAA 60

Qy      1045 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
Db        61 CAGCTGATGACCAGCGGTGTGGCTTCCAGTCGGTTTAGTGAAGTCCAGGAGCCGGAGGCC 120

Qy      1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164
Db       121 ATTGTTGTGGCCCCAACAAGAGAGCTTATCAACCAGATATATCTGGAAGCCAGGAAGTTT 180

Qy      1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA 1224
Db       181 GCTTTTGGGACGTGTGTGAGGCCTGTTGTTGTGTATGGAGGAACTAGCACCATGTACACC 240

Qy      1225 ATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGAT 1284
Db       241 ATACGTGAAGTTTTGAAAGGCTGCAACATTTTATGTGGGACACCTGGGAGACTGCTTGAC 300

Qy      1285 ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCT 1344
Db       301 ATTATTGGAAGAGGAAAGGTTGGCCTGAATAAGCTTCGCTACCTTGTACTGGATGAGGCA 360

Qy      1345 GATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
Db       361 GACCGCATGTTGGATATGGGTTTTGAACCAGACATGCGCAGGCTGGTGGCTTCTCCAGGC 420

Qy      1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
Db       421 ATGCCTCCAAAGGAGGAGCGACAGACACTTATGTTCAGTGCCACCTACCCAGAGGACATT 480

Qy      1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
Db       481 CAACGGCTTGCTGCAGATTTCTTGAAGAGAGACTACTTGTTCCTAGTTGTTGGGCAAGTG 540

Qy      1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
Db       541 GGTGGAGCCTGCAATGATGTTGAACAGCACGTAGTTGAAGTGTCCCAGTATTCAAAGCGG 600

Qy      1585 GAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAA 1644
Db       601 GAGCAGCTGGTGGAGATCCTCAGAACTACAGGGAATGAGCGCACTATGGTGTTTGTGGAA 660

Qy      1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACA 1704
Db       661 ACCAAACGAAAGGCAGATTTTATTGCCACTTTCCTCTGTCAGGAACAGATATCAACCACA 720

Qy      1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
Db       721 AGTATTCATGGTGACCGGGAACAACGGGAGCGTGAGCAAGCGCTAGGTGACTTCCGCAGT 780
```

```
Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824

Db        781 GGGAAATGTCCTGTCCTTGTGGCCACCTCGGTGGCAGCCAGAGGTCTTGACATTGAGCAT 840

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884

Db        841 GTCCAGCATGTGGTGAATTTCGACCTCCCGTCCACCATTGATGAATATGTGCACCGGATT 900

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCG 1944

Db        901 GGGAGGACAGGCCGCTGTGGGAACACTGGGAAGGCCATTTCTTTCTTTGATCCAGAGGCT 960

Qy       1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004

Db        961 GATGCCCCACTTGCTCGCTCTCTGGTTAAGATTCTTTCAGGGGCCCAGCAAGAAGTTCCC 1020

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTAC 2033

Db       1021 AAGTGGCTTGAAGAAATTGCATTTAGTGC 1049


RESULT 15
AB063484
LOCUS       AB063484                2187 bp    mRNA    linear   VRT 22-JUN-2001
DEFINITION  Oryzias latipes olvas mRNA for VASA, complete cds.
ACCESSION   AB063484
VERSION     AB063484.1  GI:14522856
KEYWORDS    .
SOURCE      Oryzias latipes cDNA to mRNA.
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1  (sites)
  AUTHORS   Shinomiya,A., Tanaka,M., Kobayashi,T., Nagahama,Y. and Hamaguchi,S.
  TITLE     The vasa-like gene, olvas, identifies the migration path of
            primordial germ cells during embryonic body formation stage in the
            medaka, Oryzias latipes
  JOURNAL   Development 42, 317-326 (2000)
REFERENCE   2  (sites)
  AUTHORS   Tanaka,M., Kinoshita,M., Kobayashi,D. and Nagahama,Y.
  TITLE     Establishment of medaka (Oryzias latipes) transgenic lines with the
            expression of green fluorescent protein fluorescence exclusively in
            germ cells: a useful model to monitor germ cells in a live
            vertebrate
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (5), 2544-2549 (2001)
  MEDLINE   21126920
   PUBMED   11226275
REFERENCE   3  (bases 1 to 2187)
  AUTHORS   Tanaka,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2001) Minoru Tanaka, Hokkaido university,
            Graduate school of Science, division of Biological science; N10,
            W8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan
            (E-mail:mtanaka@sci.hokudai.ac.jp, Tel:81-11-706-4460(ex.4460),
            Fax:81-11-706-4460)
FEATURES             Location/Qualifiers
     source          1..2187
                     /organism="Oryzias latipes"
                     /db_xref="taxon:8090"
                     /note="Japanese killifish, medaka-himedaka"
     gene            1..2187
                     /gene="olvas"
     CDS             94..1947
                     /gene="olvas"
                     /codon_start=1
                     /product="VASA"
                     /protein_id="BAB61047.1"
                     /db_xref="GI:14522857"
                     /translation="MDDWEEEETAPSFAPVSSTDAAPQRSSWNGGARDSGNDGDSWNR
                     SNRGRGGSAGRGGRGGRGRGFGRSDQDELNGGGGDSENGFRGRGRGGRGGFRSGGGER
                     GRGGGYRGRDEDVFAAGDGRGAENSDAADPERPKVTYIPPSLPEDEDSIFSHYKMGIN
                     FDKYDDILVDVSGTNLPAAIMTFEEAKLCESLENNISRSGYVKPTPVQKYGLPIISAG
                     RDLMACAQTGSGKTAAFLLPILQQLMADGVAASRFSEIQEPEAVIVAPTRELINQIYQ
                     EARKFSFGTCVRPVVVYGGVNTGYQMREIEKGCNVLCGTPGRLLDMIGRGKVGLSKVR
                     HLVLDEADRMLDMGFEPDMRRLVGSPGMPSKEERQTLMFSATFPEDIQRLAADFLKVD
                     YLFVAVGVVGGACTDVEQTFLQVTKFNKREQLLDLLRTIGSERTMVFVETKRQADFIA
                     AFLCQEKVPTTSIHGDREQREREKALADFRSGKCPVLVATSVASRGLDIPDVQHVVNF
                     DLPNTIDDYVHRIGRTGRCGNTGRAVSFYDPDVDSQLARSLVGILAKAQQEVPSWLEE
                     SAFGAHGSAAFNPSGRTFASTDSRKGGSFQDSSVKTQPAAPPAAADEDDWE"
BASE COUNT      489 a    610 c    715 g    373 t
ORIGIN

  Query Match             25.5%;  Score 552.8;  DB 5;  Length 2187;
  Best Local Similarity   61.8%;  Pred. No. 6.7e-120;
  Matches  897;  Conservative    0;  Mismatches  552;  Indels    3;  Gaps    1;

Qy        711 TGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTC 770

Db        483 TGACCCAGAACGGCCCAAAGTGACCTACATCCCCCCGAGCCTCCCGGAGGACGAGGACTC 542

Qy        771 CATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGA 830

Db        543 CATCTTCTCCCACTACAAGATGGGCATCAACTTCGACAAGTACGACGACATCCTGGTGGA 602

Qy        831 AGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTGAAGAAGCTAATCTCTGTCA 890

Db        603 CGTCAGCGGGACCAACCTGCCCGCCGCCATCATGACCTTTGAGGAGGCCAAGCTGTGCGA 662

Qy        891 GACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATA 950

Db        663 GTCGTTGGAGAACAACATCAGCAGGTCCGGATACGTCAAGCCGACGCCCGTCCAGAAGTA 722

Qy        951 CAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGG 1010

Db        723 CGGCCTCCCCATCATCTCCGCCGGCAGAGACCTGATGGCGTGCGCTCAGACCGGCTCCGG 782

Qy       1011 GAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGC 1070

Db        783 AAAAACTGCGGCGTTCCTGCTGCCCATCCTGCAGCAGCTGATGGCAGACGGCGTGGCGGC 842

Qy       1071 CAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATT 1130

Db        843 CAGCCGCTTCAGCGAGATCCAGGAGCCAGAGGCCGTGATTGTGGCTCCAACCAGGGAGCT 902

Qy       1131 GGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGT 1190

Db        903 CATCAACCAGATTTACCAGGAGGCCCGGAAGTTCTCCTTTGGGACCTGCGTGCGGCCCGT 962

Qy       1191 TGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAA 1250

Db        963 GGTGGTCTACGGAGGCGTGAACACCGGCTACCAGATGAGGGAGATCGAGAAGGGCTGCAA 1022

Qy       1251 TATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCT 1310

Db       1023 CGTGCTGTGCGGGACTCCGGGACGCCTGCTGGACATGATCGGCAGAGGAAAGGTGGGCCT 1082

Qy       1311 CAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGG 1370

Db       1083 GAGTAAAGTTCGCCACCTGGTCCTGGACGAAGCCGACCGCATGTTGGACATGGGCTTCGA 1142

Qy       1371 TCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAAC 1430

Db       1143 GCCGGACATGCGCCGGCTGGTGGGCTCCCCGGGTATGCCGTCCAAAGAGGAGCGGCAGAC 1202

Qy       1431 CCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAA 1490

Db       1203 GCTGATGTTCAGCGCCACCTTCCCCGAGGACATCCAAAGGTTGGCGGCCGACTTCCTGAA 1262

Qy       1491 GTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCA 1550

Db       1263 GGTGGACTACCTGTTTGTGGCCGTGGGGGTGGTGGGCGGAGCCTGCACCGACGTGGAGCA 1322

Qy       1551 GACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAA 1610
```

```
Db      1323 GACCTTCCTACAGGTCACCAAATTCAACAAGCGGGAGCAGCTCCTGGACCTCCTCAGGAC 1382

Qy      1611 CATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGC 1670

Db      1383 CATCGGGTCTGAGCGGACCATGGTGTTCGTGGAGACCAAGAGGCAGGCCGACTTCATCGC 1442

Qy      1671 AACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAG 1730

Db      1443 CGCCTTCCTGTGCCAGGAGAAGGTTCCGACCACCAGCATCCACGGCGACAGAGAGCAGCG 1502

Qy      1731 AGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTAC 1790

Db      1503 CGAGCGGGAGAAGGCGCTGGCGGATTTCCGCTCAGGCAAGTGTCCCGTCCTGGTGGCCAC 1562

Qy      1791 TTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCT 1850

Db      1563 CTCGGTGGCGTCCCGCGGCCTGGACATCCCTGACGTCCAGCACGTGGTGAACTTTGACCT 1622

Qy      1851 TCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATAC 1910

Db      1623 GCCCAACACCATTGACGACTACGTCCACCGCATCGGGAGAACGGGCCGCTGCGGCAACAC 1682

Qy      1911 TGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGT 1970

Db      1683 CGGCAGGGCCGTGTCCTTCTACGACCCCGACGTGGACAGTCAGCTGGCCCGCTCGCTGGT 1742

Qy      1971 AAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAG 2030

Db      1743 CGGCATCCTGGCCAAGGCTCAGCAGGAGGTGCCGTCATGGCTGGAGGAGTCGGCGTTCGG 1802

Qy      2031 TACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATAC 2090

Db      1803 CGCTCACGGTTCCGCCGCC---TTCAACCCTTCGGGAAGGACGTTTGCTTCCACAGACTC 1859

Qy      2091 CAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCC 2150

Db      1860 CAGGAAGGGTGGTTCTTTCCAGGACAGCAGTGTGAAGACGCAGCCTGCAGCGCCCCCTGC 1919

Qy      2151 AGTAGATGATGA 2162

Db      1920 TGCTGCTGATGA 1931
```

Search completed: June 10, 2003, 09:43:34
Job time : 5469 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:         June 10, 2003, 14:53:30 ; Search time 28 Seconds
                                          (without alignments)
                                          6434.756 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  3954
Sequence:       1 atgggggatgaagattggga.........tagatgatgagtcatgggat 2172

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters:      225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111758_26006/app_query.fasta_1.2311
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_23_@runat_05062003_111758_26006 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      SwissProt_40:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3813	96.4	724	1	DDX4_HUMAN	Q9nqi0 homo sapien
2	3249	82.2	702	1	DDX4_MOUSE	Q61496 mus musculu
3	3247.5	82.1	713	1	DDX4_RAT	Q64060 rattus norv
4	1290	32.6	661	1	VASA_DROME	P09052 drosophila
5	1229	31.1	661	1	DDX3_MOUSE	Q62167 mus musculu
6	1224.5	31.0	697	1	AN3_XENLA	P24346 xenopus lae
7	1215	30.7	661	1	DDX3_HUMAN	O00571 homo sapien
8	1197	30.3	660	1	PL10_MOUSE	P16381 mus musculu
9	1182.5	29.9	660	1	DDXY_HUMAN	O15523 homo sapien
10	1126.5	28.5	604	1	DED1_YEAST	P06634 saccharomyc
11	1109	28.0	636	1	DED1_SCHPO	O13370 schizosacch
12	1108	28.0	617	1	DBP1_YEAST	P24784 saccharomyc
13	1096.5	27.7	763	1	GLH1_CAEEL	P34689 caenorhabdi
14	992.5	25.1	720	1	GLH3_CAEEL	O01836 caenorhabdi
15	820.5	20.8	550	1	DBP2_SCHPO	P24782 schizosacch
16	815	20.6	546	1	DBP2_YEAST	P24783 saccharomyc
17	795.5	20.1	575	1	RM62_DROME	P19109 drosophila
18	777.5	19.7	614	1	DDX5_HUMAN	P17844 homo sapien
19	777.5	19.7	614	1	DDX5_MOUSE	Q61656 mus musculu
20	771.5	19.5	1156	1	GLH4_CAEEL	O76743 caenorhabdi
21	754	19.1	607	1	DB10_NICSY	P46942 nicotiana s
22	754	19.1	650	1	DD17_HUMAN	Q92841 homo sapien
23	748	18.9	568	1	HE64_TRYBB	Q26696 trypanosoma
24	693.5	17.5	523	1	DBP3_YEAST	P20447 saccharomyc
25	692.5	17.5	578	1	YBX6_SCHPO	Q10202 schizosacch
26	685	17.3	454	1	RHLE_ECOLI	P25888 escherichia
27	684	17.3	245	1	DEA2_MOUSE	Q62095 mus musculu
28	675	17.1	622	1	ABS_HUMAN	Q9ujv9 homo sapien
29	668.5	16.9	619	1	ABS_DROME	Q9v3c0 drosophila
30	621	15.7	524	1	EXP9_STRPN	P35599 streptococc
31	594	15.0	543	1	RRP3_YEAST	P38712 saccharomyc
32	587	14.8	588	1	PR28_YEAST	P23394 saccharomyc
33	575.5	14.6	481	1	YA88_SCHPO	Q09775 schizosacch
34	569	14.4	563	1	DEAD_MYCTU	Q11039 mycobacteri
35	567.5	14.4	642	1	DEAD_KLEPN	P33906 klebsiella
36	566.5	14.3	628	1	DEAD_ECOLI	P23304 escherichia
37	563.5	14.3	851	1	DD21_MOUSE	Q9jik5 mus musculu
38	562.5	14.2	613	1	DEAD_HAEIN	P44586 haemophilus
39	560	14.2	715	1	DD21_HUMAN	Q9nr30 homo sapien
40	556	14.1	601	1	DEAD_BUCAI	P57453 buchnera ap
41	554.5	14.0	848	1	YA47_SCHPO	Q09719 schizosacch
42	548.5	13.9	564	1	ROK1_YEAST	P45818 saccharomyc
43	546.5	13.8	479	1	DBPA_BACSU	P42305 bacillus su
44	542	13.7	457	1	DBPA_ECOLI	P21693 escherichia
45	533	13.5	754	1	YAJ3_SCHPO	Q09903 schizosacch

ALIGNMENTS

```
RESULT 1
DDX4_HUMAN
ID   DDX4_HUMAN     STANDARD;      PRT;   724 AA.
AC   Q9NQI0; Q9NYB1; Q9NT92;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 4 (VASA homolog).
GN   DDX4 OR VASA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC   TISSUE=Testis;
RX   MEDLINE=20402578; PubMed=10920202;
RA   Castrillon D.H., Quade B.J., Wang T.Y., Quigley C., Crum C.P.;
RT   "The human VASA gene is specifically expressed in the germ cell
RT   lineage.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:9585-9590(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Rocha D., Affara N.;
RT   "Cloning and characterization of the human VASA gene.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 90-724 FROM N.A.
RC   TISSUE=Testis;
RA   Blum H., Bauersachs S., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May play a role in germ cell development.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Expressed only in ovary and testis. Expressed
CC       in migratory primordial germ cells in the region of the gonadal
CC       ridge in both sexes.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC       DDX4/VASA SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
```

```
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AY004154; AAF86585.1; -.
DR   EMBL; AF262962; AAF72705.1; -.
DR   EMBL; AL137462; CAB70750.1; -.
DR   HSSP; Q58083; 1HV8.
DR   Genew; HGNC:18700; DDX4.
DR   MIM; 605281; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 2.
DR   Pfam; PF00271; helicase_C; 2.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.
FT   NP_BIND      332    339       ATP (POTENTIAL).
FT   SITE         446    449       DEAD BOX.
FT   DOMAIN        58    234       GLY-RICH.
FT   CONFLICT     556    556       I -> T (IN REF. 1).
FT   CONFLICT     713    713       Q -> R (IN REF. 1).
SQ   SEQUENCE   724 AA;  79307 MW;  6D313DD98B177067 CRC64;

Alignment Scores:
Pred. No.:                 1.57e-261      Length:          724
Score:                     3813.00        Matches:         722
Percent Similarity:        99.86%         Conservative:    1
Best Local Similarity:     99.72%         Mismatches:      1
Query Match:               96.43%         Indels:          0
DB:                        1              Gaps:            0

US-09-714-865-15 (1-2172) x DDX4_HUMAN (1-724)

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAspGluAspTrpGluAlaGluIleAsnProHisMetSerSerTyrValProIle 20

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 PheGluLysAspArgTyrSerGlyGluAsnGlyAspAsnPheAsnArgThrProAlaSer 40

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerSerGluMetAspAspGlyProSerArgArgAspHisPheMetLysSerGlyPheAla 60

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerGlyArgAsnPheGlyAsnArgAspAlaGlyGluCysAsnLysArgAspAsnThrSer 80

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrMetGlyGlyPheGlyValGlyLysSerPheGlyAsnArgGlyPheSerAsnSerArg 100

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheGluAspGlyAspSerSerGlyPheTrpArgGluSerSerAsnAspCysGluAspAsn 120

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ProThrArgAsnArgGlyPheSerLysArgGlyGlyTyrArgAspGlyAsnAsnSerGlu 140

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AlaSerGlyProTyrArgArgGlyGlyArgGlySerPheArgGlyCysArgGlyGlyPhe 160

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GlyLeuGlySerProAsnAsnAspLeuAspProAspGluCysMetGlnArgThrGlyGly 180

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LeuPheGlySerArgArgProValLeuSerGlyThrGlyAsnGlyAspThrSerGlnSer 200

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ArgSerGlySerGlySerGluArgGlyGlyTyrLysGlyLeuAsnGluGluValIleThr 220

Qy        661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyGluSerSerAspThrGln 240

Qy        721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePheAla 260

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSerGly 280

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 HisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeuAsn 300

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIlePro 320

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 IleIleLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 340

Qy       1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaPheLeuLeuProIleLeuAlaHisMetMetHisAspGlyIleThrAlaSerArgPhe 360

Qy       1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||.|||||||
Db        361 LysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuValAsnGln 380

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaValValIleTyr 400

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyGlyThrGlnLeuGlyHisSerIleArgGlnIleValGlnGlyCysAsnIleLeuCys 420

Qy       1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLysGlnIle 440

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlyProGluMet 460

Qy       1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 LysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeuMetPhe 480

Qy       1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SerAlaThrPheProGluGluIleGlnArgLeuAlaAlaGluPheLeuLysSerAsnTyr 500

Qy       1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThrValLeu 520

Qy       1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GlnValGlyGlnPheSerLysArgGluLysLeuValGluIleLeuArgAsnIleGlyAsp 540

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db       541 GluArgThrMetValPheValGluThrLysLysLysAlaAspPheIleAlaThrPheLeu 560

Qy      1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       561 CysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGluArgGlu 580

Qy      1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       581 GlnAlaLeuGlyAspPheArgPheGlyLysCysProValLeuValAlaThrSerValAla 600

Qy      1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 AlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAspLeuProSerThr 620

Qy      1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       621 IleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAla 640

Qy      1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       641 IleSerPhePheAspLeuGluSerAspAsnHisLeuAlaGlnProLeuValLysValLeu 660

Qy      1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 ThrAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerThrTyrIle 680

Qy      2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       681 ProGlyPheSerGlySerThrArgGlyAsnValPheAlaSerValAspThrArgLysGly 700

Qy      2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
             |||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db       701 LysSerThrLeuAsnThrAlaGlyPheSerSerSerGlnAlaProAsnProValAspAsp 720

Qy      2161 GAGTCATGGGAT 2172
             ||||||||||||
Db       721 GluSerTrpAsp 724

RESULT 2
DDX4_MOUSE
ID   DDX4_MOUSE     STANDARD;      PRT;   702 AA.
AC   Q61496; Q9D5X7;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 4 (VASA homolog) (Mvh).
GN   DDX4 OR VASA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE OF 60-702 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Testis;
RX   MEDLINE=95083681; PubMed=7991615;
RA   Fujiwara Y., Komiya T., Kawabata H., Sato M., Fujimoto H.,
RA   Furusawa M., Noce T.;
RT   "Isolation of a DEAD-family protein gene that encodes a murine
RT   homolog of Drosophila vasa and its specific expression in germ cell
RT   lineage.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:12258-12262(1994).
CC   -!- FUNCTION: May play a role in germ cell development.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND PERINUCLEAR.
CC   -!- TISSUE SPECIFICITY: Testis.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN SPERMATOGENIC CELLS FROM THE
CC       SPERMATOCYTE STAGE TO THE ROUND SPERMATID STAGE.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC       DDX4/VASA SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AK014844; BAB29578.1; -.
DR   EMBL; D14859; BAA03584.1; -.
DR   HSSP; Q58083; 1HV8.
DR   MGD; MGI:102670; Ddx4.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 2.
DR   Pfam; PF00271; helicase_C; 2.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.
FT   NP_BIND     305    312       ATP (POTENTIAL).
FT   SITE        419    422       DEAD BOX.
FT   DOMAIN       58    207       GLY-RICH.
FT   CONFLICT    152    152       R -> C (IN REF. 2).
FT   CONFLICT    156    160       LFGSR -> FLVLG (IN REF. 2).
FT   CONFLICT    277    277       A -> R (IN REF. 2).
FT   CONFLICT    291    291       S -> T (IN REF. 2).
FT   CONFLICT    368    369       RA -> IS (IN REF. 2).
FT   CONFLICT    423    423       R -> S (IN REF. 2).
FT   CONFLICT    430    430       G -> A (IN REF. 2).
FT   CONFLICT    433    433       M -> I (IN REF. 2).
FT   CONFLICT    448    448       R -> H (IN REF. 2).
FT   CONFLICT    472    472       S -> N (IN REF. 2).
FT   CONFLICT    499    509       SKREKLVEILR -> QKEKSLLRFYE (IN REF. 2).
FT   CONFLICT    540    540       T -> S (IN REF. 2).
FT   CONFLICT    680    702       HTLNTAGISSSQAPNPVDDESWD -> AHVEYSGDFFFTSS
FT                                QSS (IN REF. 2).
SQ   SEQUENCE   702 AA;  76470 MW;  5D1E4CC0BDF39748 CRC64;

Alignment Scores:
Pred. No.:               1.01e-221      Length:         702
Score:                   3249.00        Matches:        626
Percent Similarity:      90.42%         Conservative:   35
Best Local Similarity:   85.64%         Mismatches:     34
Query Match:             82.17%         Indels:         36
DB:                      1              Gaps:           7

US-09-714-865-15 (1-2172) x DDX4_MOUSE (1-702)

Qy         1 ATGGGGGATGAAGATTGGGAAGCAGAAATC---AACCCTCATATGTCTTCCTATGTTCCC 57
             ||||||||||||||||||||||||||||||      |||||:::||||||||||||||
```

```
Db          1 MetGlyAspGluAspTrpGluAlaGluIleLeuLysProHisValSerSerTyrValPro 20

Qy         58 ATATTTGAGAAGGATAGGTAT---TCTGGAGAAAATGGAGACAATTTTAACAGGACTCCA 114
              :::||||||||||:::|||   ||||||   |||||||||   |||||||||||
Db         21 ValPheGluLysAspLysTyrSerSerGlyAlaAsnGlyAspThrPheAsnArgThrSer 40

Qy        115 GCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGA 174
              |||||||||
Db         41 AlaSerSer--------------------------------------------------- 43

Qy        175 TTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT 234
                                               |||   ||||||    :::|||::::::|||
Db         44 ------------------------------AspIleGlyGluSerSerLysLysGluAsn 53

Qy        235 ACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAAC 294
              |||||||||   ||||||||||||   ||||||   ||||||||||||||||||    |||
Db         54 ThrSerThrThrGlyGlyPheGlyArgGlyLysGlyPheGlyAsnArgGlyPheLeuAsn 73

Qy        295 AGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAA 354
              ::::::||||||:::||||||||||||||||||||:::||||||:::||||||||||||
Db         74 AsnLysPheGluGluGlyAspSerSerGlyPheTrpLysGluSerAsnAsnAspCysGlu 93

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
              ||||||   ||||||:::||||||||||||||||||||||||   :::|||||||||:::
Db         94 AspAsnGlnThrArgSerArgGlyPheSerLysArgGlyGlyCysGlnAspGlyAsnAsp 113

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGA 474
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        114 SerGluAlaSerGlyProPheArgArgGlyGlyArgGlySerPheArgGlyCysArgGly 133

Qy        475 GGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACT 534
              ||||||||||||||||   ||||||:::::::   |||   |||:::      ||||||
Db        134 GlyPheGlyLeuGlyArgProAsnSerGluSerAspGlnAspGlnGlyThrGlnArgGly 153

Qy        535 GGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCT 594
              |||||||||||||||||||:::|||         |||   :::|||||||||||||||
Db        154 GlyGlyLeuPheGlySerArgLysProAlaAlaSerAspSerGlyAsnGlyAspThrTyr 173

Qy        595 CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654
              |||||||||||||||||||||        ||||||||||||||||||||||||||||||||
Db        174 GlnSerArgSerGlySerGly------ArgGlyGlyTyrLysGlyLeuAsnGluGluVal 191

Qy        655 ATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGAT 714
              :::||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db        192 ValThrGlySerGlyLysAsnSerTrpLysSerGluThrGluGlyGlyGluSerSerAsp 211

Qy        715 ACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATC 774
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 SerGlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIle 231

Qy        775 TTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 PheAlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluVal 251

Qy        835 TCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACA 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 SerGlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThr 271

Qy        895 CTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGT 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 LeuAsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSer 291

Qy        955 ATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAG 1014
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 IleProIleValLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLys 311

Qy       1015 ACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGT 1074
              ||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        312 ThrAlaAlaPheLeuLeuProIleLeuAlaHisMetMetArgAspGlyIleThrAlaSer 331

Qy       1075 CGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTC 1134
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        332 ArgPheLysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuIle 351

Qy       1135 AACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTT 1194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 AsnGlnIleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaValVal 371

Qy       1195 ATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATA 1254
              ||||||||||||||||   |||||||||:::|||||||||||||||||||||||||||
Db        372 IleTyrGlyGlyThrGlnPheGlyHisSerValArgGlnIleValGlnGlyCysAsnIle 391

Qy       1255 TTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAA 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 LeuCysAlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLys 411

Qy       1315 CAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCA 1374
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 GlnValLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlyPro 431

Qy       1375 GAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTT 1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 GluMetLysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeu 451

Qy       1435 ATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCA 1494
              :::|||||||||||||||||||||||||||||||||||||||   :::|||||||||||
Db        452 LeuPheSerAlaThrPheProGluGluIleGlnArgLeuAlaGlyAspPheLeuLysSer 471

Qy       1495 AATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACC 1554
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SerTyrLeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThr 491

Qy       1555 GTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATA 1614
              :::|||||||||||||:::|||||||||||||||||||||||||||||||||||||||||
Db        492 IleLeuGlnValGlyGlnTyrSerLysArgGluLysLeuValGluIleLeuArgAsnIle 511

Qy       1615 GGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        512 GlyAspGluArgThrMetValPheValGluThrLysLysLysAlaAspPheIleAlaThr 531

Qy       1675 TTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 PheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGlu 551

Qy       1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db        552 ArgGluGlnAlaLeuGlyAspPheArgCysGlyLysCysProValLeuValAlaThrSer 571

Qy       1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 ValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAspLeuPro 591

Qy       1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 SerThrIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGly 611

Qy       1915 AGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAA 1974
              |||||||||||||||||||   :::|||||||||||||||||||||||||||||||||||
Db        612 ArgAlaIleSerPhePheAspThrAspSerAspAsnHisLeuAlaGlnProLeuValLys 631

Qy       1975 GTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACA 2034
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        632 ValLeuSerAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerThr 651

Qy       2035 TACATT---CCTGGCTTCAGTGGTAGTACAAGAGGAAAC---GTGTTTGCATCAGTTGAT 2088
              |||:::   |||   ||||||   |||||||||||       |||||||||||||||||
Db        652 TyrValProProSerPheSerSerSerThrArgGlyGlyAlaValPheAlaSerValAsp 671

Qy       2089 ACCAGAAAG---------GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
              |||||||||         ||||||   |||||||||||||||||||   ||||||||:::
Db        672 ThrArgLysAsnTyrGlnGlyLysHisThrLeuAsnThrAlaGlyIleSerSerSerGln 691

Qy       2140 GCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
              |||||||||||||||||||||||||||||||||
Db        692 AlaProAsnProValAspAspGluSerTrpAsp 702
```

```
RESULT 3
DDX4_RAT
ID   DDX4_RAT       STANDARD;      PRT;   713 AA.
AC   Q64060;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 4 (VASA homolog) (rVLG).
GN   DDX4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Testis;
RX   MEDLINE=95160706; PubMed=7857296;
RA   Komiya T., Tanigawa Y.;
RT   "Cloning of a gene of the DEAD box protein family which is
RT   specifically expressed in germ cells in rats.";
RL   Biochem. Biophys. Res. Commun. 207:405-410(1995).
CC   -!- FUNCTION: May play a role in germ cell development.
CC   -!- TISSUE SPECIFICITY: Testis.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC       DDX4/VASA SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; S75275; AAB33364.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.
FT   NP_BIND     317    324       ATP (POTENTIAL).
FT   SITE        431    434       DEAD BOX.
FT   DOMAIN       58    219       GLY-RICH.
SQ   SEQUENCE   713 AA;  77955 MW;  A31A4542EF7237F6 CRC64;

Alignment Scores:
Pred. No.:              1.29e-221     Length:         713
Score:                  3247.50       Matches:        625
Percent Similarity:     90.55%        Conservative:   36
Best Local Similarity:  85.62%        Mismatches:     46
Query Match:            82.13%        Indels:         23
DB:                     1             Gaps:           6

US-09-714-865-15 (1-2172) x DDX4_RAT (1-713)

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATC---AACCCTCATATGTCTTCCTATGTTCCC 57
                ||||||||||||||||||||||||||||     |||||:::||||||||||||||
Db          1 MetGlyAspGluAspTrpGluAlaGluIleLeuLysProHisValSerSerTyrValPro 20

Qy         58 ATATTTGAGAAGGATAGGTAT---TCTGGAGAAAATGGAGACAATTTTAACAGGACTCCA 114
                :::|||||||||||:::|||   ||||||    |||||||||    |||||||||||
Db         21 ValPheGluLysAspLysTyrSerSerGlyAlaAsnGlyAspThrPheAsnArgThrSer 40

Qy        115 GCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGA 174
                ||||||||||||||||:::|||||||||||    |||||||||||||:::||||||
Db         41 AlaSerSerSerGluMetGluAspGlyProSerGlyArgAspHisPheMetArgSerGly 60

Qy        175 TTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT 234
                |||:::|||||||||||·  |||||||||||    ||||||    :::||||||:::
Db         61 PheSerSerGlyArgAsnLeuGlyAsnArgAspIleGlyGluSerSerLysArgGluThr 80

Qy        235 ACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAAC 294
                ||||||||  ||||||||||||   ||||||   ||||||||||||||||    |||
Db         81 ThrSerThrThrGlyGlyPheGlyArgGlyLysGlyPheGlyAsnArgGlyPheLeuAsn 100

Qy        295 AGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAA 354
                ::::::||||||:::||||||||||||||||||:::||||||:::||||||||||||
Db        101 AsnLysPheGluGluGlyAspSerSerGlyPheTrpLysGluSerThrAsnAspCysGlu 120

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
                |||        ||||||:::||||||||||||||||||||||||    |||||||||:::
Db        121 AspThrGlnThrArgSerArgGlyPheSerLysArgGlyGlyTyrProAspGlyAsnAsp 140

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGA 474
                ||||||||||||||||:::||||||||||||||    |||
Db        141 SerGluAlaSerGlyProPheArgArgGlyGlyArgAspSer------------------ 154

Qy        475 GGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACT 534
                                           :::   |||   |||:::      ||||||
Db        155 ---------------------------GluTyrAspGlnAspGlnGlySerGlnArgGly 165

Qy        535 GGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCT 594
                ||||||||||||||||||:::|||      |||     :::|||:::|||||||||
Db        166 GlyGlyLeuPheGlySerArgLysProAlaAlaSerAspSerGlySerGlyAspThrPhe 185

Qy        595 CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTAAATGAAGAAGTA 654
                |||||||||||||||:::          ||||||   ||||||||||||||||||||
Db        186 GlnSerArgSerGlyAsnAla------ArgGlyAlaTyrLysGlyLeuAsnGluGluVal 203

Qy        655 ATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGAT 714
                :::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValThrGlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyGluSerSerAsp 223

Qy        715 ACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATC 774
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 IleGlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIle 243

Qy        775 TTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTG 834
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 PheAlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluVal 263

Qy        835 TCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACA 894
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 SerGlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThr 283

Qy        895 CTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGT 954
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 LeuAsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSer 303

Qy        955 ATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAG 1014
                |||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        304 IleProIleValLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLys 323

Qy       1015 ACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGT 1074
                |||||||||||||||||||||||||||||||||||    ||||||||||||||||
Db        324 ThrAlaAlaPheLeuLeuProIleLeuAlaHisMetMetArgAspGlyIleThrAlaSer 343

Qy       1075 CGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTC 1134
                |||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        344 ArgPheLysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuIle 363

Qy       1135 AACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTT 1194
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 AsnGlnIleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaValVal 383

Qy       1195 ATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATA 1254
                ||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db        384 IleTyrGlyGlyThrGlnPheGlyHisSerIleArgGlnIleValGlnGlyCysAsnIle 403

Qy       1255 TTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAA 1314
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        404 LeuCysAlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLys 423

Qy       1315 CAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCA 1374
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GlnValLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlyPro 443

Qy       1375 GAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTT 1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GluMetLysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeu 463

Qy       1435 ATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCA 1494
              :::|||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        464 LeuPheSerAlaThrPheProGluGluIleGlnArgLeuAlaGlyGluPheLeuLysSer 483

Qy       1495 AATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACC 1554
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        484 AsnTyrLeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnSer 503

Qy       1555 GTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATA 1614
              :::|||||||||||            |||:::   |||||||||||||||||||||||||
Db        504 IleLeuGlnValGlyProValPheLysLysArgLysLeuValGluIleLeuArgAsnIle 523

Qy       1615 GGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
              ||||||||||||   ||||||||||||||||||||||||||||||||||   ||||||
Db        524 GlyAspGluArgProMetValPheValGluThrLysLysLysAlaAspPheIleAlaThr 543

Qy       1675 TTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 PheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGlu 563

Qy       1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
              ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        564 ArgGluGlnAlaLeuGlyAspPheArgCysGlyLysCysProValLeuValAlaThrSer 583

Qy       1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        584 ValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAsnLeuPro 603

Qy       1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 SerThrIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGly 623

Qy       1915 AGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAA 1974
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        624 ArgAlaIleSerPhePheAspThrGluSerAspAsnHisLeuAlaGlnProLeuValLys 643

Qy       1975 GTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACA 2034
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        644 ValLeuSerAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerSer 663

Qy       2035 TACATT---CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACC 2091
              |||      |||   ||||||   ||||||||||||   |||||||||||   ||||||
Db        664 TyrAlaProProSerPheSerAsnSerThrArgGlyAlaValPheAlaSerPheAspThr 683

Qy       2092 AGAAAG---------GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
              ||||||         ||||||:::||||||||||||||||||   ||||||::::::|||
Db        684 ArgLysAsnPheGlnGlyLysAsnThrLeuAsnThrAlaGlyIleSerSerAlaGlnAla 703

Qy       2143 CCCAATCCAGTAGATGATGAGTCATGGGAT 2172
              ||||||||||||||||||||||||||||||
Db        704 ProAsnProValAspAspGluSerTrpAsp 713
```

RESULT 4

```
VASA_DROME
ID   VASA_DROME     STANDARD;      PRT;   661 AA.
AC   P09052; Q24582; Q9V3Q8;
DT   01-NOV-1988 (Rel. 09, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Vasa protein (Antigen Mab46F11).
GN   VAS OR BG:DS00929.14 OR CG3506.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89014721; PubMed=3140040;
RA   Lasko P.F., Ashburner M.;
RT   "The product of the Drosophila gene vasa is very similar to
RT   eukaryotic initiation factor-4A.";
RL   Nature 335:611-617(1988).
RN   [2]
RP   REVISIONS.
RA   Lasko P.F.;
RL   Submitted (DEC-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89028669; PubMed=3052853;
RA   Hay B., Jan L.Y., Jan Y.N.;
RT   "A protein component of Drosophila polar granules is encoded by vasa
RT   and has extensive sequence similarity to ATP-dependent helicases.";
RL   Cell 55:577-587(1988).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=99403001; PubMed=10471707;
RA   Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA   Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA   Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA   Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA   Celniker S., Rubin G.M.;
RT   "An exploration of the sequence of a 2.9-Mb region of the genome of
RT   Drosophila melanogaster: the Adh region.";
RL   Genetics 153:179-219(1999).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
```

```
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: THE VASA PROTEIN IS REQUIRED ONLY IN THE FEMALE GERM
CC       LINE. IT IS IMPORTANT FOR OOCYTE FORMATION AND IN THE
CC       SPECIFICATION OF THE POSTERIOR STRUCTURES OF THE EMBRYO.
CC   -!- DEVELOPMENTAL STAGE: MATERNALLY EXPRESSED (DURING OOGENESIS).
CC       FUNCTION DURING EARLY EMBRYOGENESIS.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC       DDX4/VASA SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X12945; CAA31405.1; -.
DR   EMBL; X12946; CAA31405.1; JOINED.
DR   EMBL; M23560; AAA29013.1; -.
DR   EMBL; AE003412; AAF44917.1; -.
DR   EMBL; AE003646; AAF53438.1; -.
DR   PIR; A31922; A31922.
DR   PIR; S01676; S01676.
DR   HSSP; Q58083; 1HV8.
DR   FlyBase; FBgn0003970; vas.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.
FT   NP_BIND     289    296       ATP (BY SIMILARITY).
FT   SITE        399    402       DEAD BOX.
FT   DOMAIN       93    127       5 X 7 AA TANDEM REPEATS OF [FS]-R-G-G-
FT                                [EQ]-G-G.
FT   REPEAT       93     99       1.
FT   REPEAT      100    106       2.
FT   REPEAT      107    113       3.
FT   REPEAT      114    120       4.
FT   REPEAT      121    127       5.
FT   CONFLICT     35     35       A -> R (IN REF. 3).
FT   CONFLICT    153    165       MISSING (IN REF. 3).
FT   CONFLICT    192    192       V -> A (IN REF. 1 AND 3).
FT   CONFLICT    265    265       Y -> F (IN REF. 1).
FT   CONFLICT    322    322       V -> C (IN REF. 3).
FT   CONFLICT    452    452       F -> S (IN REF. 1).
FT   CONFLICT    582    582       R -> C (IN REF. 1).
FT   CONFLICT    594    594       D -> H (IN REF. 3).
SQ   SEQUENCE   661 AA;  72331 MW;  8617C25CCB3130B9 CRC64;

Alignment Scores:
Pred. No.:              1.9e-83      Length:        661
Score:                  1290.00      Matches:       314
Percent Similarity:     56.46%       Conservative:  101
Best Local Similarity:  42.72%       Mismatches:    230
Query Match:            32.63%       Indels:        90
DB:                     1            Gaps:          24

US-09-714-865-15 (1-2172) x VASA_DROME (1-661)

Qy         10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG--- 66
              :::||||||:::    |||                           |||||    :::
Db          3 AspAspTrpAspAspGlu---------------------------ProIleValAspThr 13

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              :::            |||:::     |||:::   :::          :::   |||
Db         14 ArgGlyAlaArgGlyGlyAspTrpSerAspAspGluAspThrAlaLysSerPheSerGly 33

Qy        127 GAAATGGAT---GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCT 183
              |||   :::   ||||||                                  |||
Db         34 GluAlaGluGlyAspGlyValGlyGly---------------SerGlyGlyGluGlyGly 48

Qy        184 GGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACA 243
              |||         ||||||||||||
Db         49 GlyTyrGlnGlyGlyAsnArgAsp---------------------------------Val 57

Qy        244 ATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTT 303
                 |||      |||   |||:::       |||       |||:::
Db         58 PheGlyArgIleGlyGlyGlyArgGlyGlyGlyAlaGlyGlyTyrArgGlyGlyAsnArg 77

Qy        304 GAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCA 363
              :::   |||            |||      ||||||       |||
Db         78 AspGlyGlyGlyPheHisGlyGlyArgArgGluGlyGluArgAspPhe----------- 93

Qy        364 ACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCT 423
                       ||||||         ||||||:::|||   |||
Db         94 ---------ArgGly------GlyGluGlyGlyPheArgGlyGlyGlnGlyGlySerArg 108

Qy        424 TCAGGGCCA---TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
                 |||         ||||||||:::|||   |||||||||   |||||||||
Db        109 GlyGlyGlnGlyGlySerArgGlyGlyGlnGlyGlyPheArgGlyGlyGluGlyGlyPhe 128

Qy        481 ---GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGT 537
                 |||         |||   |||   |||               :::|||   |||
Db        129 ArgGlyArgLeuTyrGluAsnGluAspGlyAsp---------------GluArgArgGly 143

Qy        538 GGCCTT---------------TTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAAT 582
                 |||               |||   ||||||      |||
Db        144 ArgLeuAspArgGluGluArgGlyGlyGluArgArgGlyArgLeuAsp----------- 159

Qy        583 GGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGA------GGTGGTTAC--- 633
                       :::   |||   |||         ||||||      ||||||:::
Db        160 ---------ArgGluGluArgGlyGlyGlyGluArgGlyGluArgGlyAspGlyGlyPheAla 176

Qy        634 ---AAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAA 690
                 :::      ||||||:::   |||         :::
Db        177 ArgArgArgArgAsnGluAspAspIleAsnAsnAsn----------------------- 188

Qy        691 GCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCT 750
                          |||   :::      :::      |||||||||||
Db        189 ------AsnAsnIleValGluAspValGluArgLysArgGluPheTyrIleProProGlu 206

Qy        751 CCACCTGAGGATGAGGACTCCATCTTTGCACAT---TATCAGACAGGCATAAACTTCGAC 807
              |||      |||         ||||||:::            :::||||||:::|||
Db        207 ProSerAsnAspAlaIleGluIlePheSerSerGlyIleAlaSerGlyIleHisPheSer 226

Qy        808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
              ||||||:::      |||   |||:::|||:::|||   |||      |||      |||
Db        227 LysTyrAsnAsnIleProValLysValThrGlySerAspValProGlnProIleGlnHis 246

Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
              |||      |||:::|||         :::      :::|||:::   |||:::||||||
Db        247 PheThrSerAlaAspLeuArgAspIleIleIleAspAsnValAsnLysSerGlyTyrLys 266

Qy        928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
                    ||||||:::||||||   |||||||||:::|||      :::||||||||||||
Db        267 IleProThrProIleGlnLysCysSerIleProValIleSerSerGlyArgAspLeuMet 286

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        287 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerLys 306

Qy       1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
              ::::::   |||                        ||||||      |||:::   :::
Db        307 LeuLeuGluAsp------------ProHisGluLeuGluLeuGlyArgProGlnValVal 322

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTGGAAGCCAGAAAATTTTCT 1167
              ||||||:::||||||||||||||||         ||||||:::   ||||||||||||:::
Db        323 IleValSerProThrArgGluLeuAlaIleGlnIlePheAsnGluAlaArgLysPheAla 342

Qy       1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
```

```
                 |||   :::   ::::::     ::::::|||||||||||       |||
Db        343 PheGluSerTyrLeuLysIleGlyIleValTyrGlyGlyThrSerPheArgHisGlnAsn 362

Qy       1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
                    |||   :::||||||:::::::::   |||||||||||||||||:::|||
Db        363 GluCysIleThrArgGlyCysHisValValIleAlaThrProGlyArgLeuLeuAspPhe 382

Qy       1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
              :::   :::         |||        :::       ::::::::|||||||||||||||||
Db        383 ValAspArgThrPheIleThrPheGluAspThrArgPheValValLeuAspGluAlaAsp 402

Qy       1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
              ||||||||||||||||||||          :::|||:::::::::::::::       |||
Db        403 ArgMetLeuAspMetGlyPheSerGluAspMetArgArgIleMetThrHisValThrMet 422

Qy       1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
                    :::   :::   ||||||||||||||||||||||||||||||||||||||||
Db        423 ------ArgProGluHisGlnThrLeuMetPheSerAlaThrPheProGluGluIleGln 440

Qy       1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGT 1527
              |||:::|||   ||||||||||||    ||||||:::|||||||||:::|||   ||||||
Db        441 ArgMetAlaGlyGluPheLeuLys---AsnTyrValPheValAlaIleGlyIleValGly 459

Qy       1528 GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAA 1587
              |||||||||   ||||||:::||||||:::    :::|||    :::::::::||||||
Db        460 GlyAlaCysSerAspValLysGlnThrIleTyrGluValAsnLysTyrAlaLysArgSer 479

Qy       1588 AAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACT 1647
              ||||||:::|||||||||             |||      |||:::||||||||||||||||
Db        480 LysLeuIleGluIleLeuSerGluGlnAlaAsp---GlyThrIleValPheValGluThr 498

Qy       1648 AAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGT 1707
              |||:::   |||||||||    |||:::||||||   :::::::::      |||||||||
Db        499 LysArgGlyAlaAspPheLeuAlaSerPheLeuSerGluLysGluPheProThrThrSer 518

Qy       1708 ATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGA 1767
              ||||||||||||||   |||    :::|||||||||||||||    ||||||:::    |||
Db        519 IleHisGlyAspArgLeuGlnSerGlnArgGluGlnAlaLeuArgAspPheLysAsnGly 538

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
                       ||||||:::||||||||||||||||:::||||||||||||||||:::|||:::
Db        539 SerMetLysValLeuIleAlaThrSerValAlaSerArgGlyLeuAspIleLysAsnIle 558

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              :::||||||||||||:::|||:::||||||     |||:::||||||||||||||||:::|||
Db        559 LysHisValIleAsnTyrAspMetProSerLysIleAspAspTyrValHisArgIleGly 578

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||   ||||||   |||||||||   ||||||||||||    |||   |||
Db        579 ArgThrGlyArgValGlyAsnAsnGlyArgAlaThrSerPhePheAspProGluLysAsp 598

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
                    :::|||      |||||||||:::|||        :::    |||    ||||||
Db        599 ArgAlaIleAlaAlaAspLeuValLysIleLeuGluGlySerGlyGlnThrValProAsp 618

Qy       2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
              :::|||                         |||       |||      |||
Db        619 PheLeuArg---------------ThrCysGlyAlaGlyGlyAspGlyGlyTyrSerAsn 633

Qy       2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
                    |||      ||||||   |||   |||:::              |||      |||
Db        634 GlnAsnPheGlyGlyValAspValArg---GlyArgGly------AsnTyrValGly--- 649

Qy       2128 TCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
                           |||   |||        ::::::|||   ||||||
Db        650 ---------AspAlaThrAsnValGluGluGluGluGlnTrpAsp 661

RESULT 5
DDX3_MOUSE
ID   DDX3_MOUSE     STANDARD;      PRT;   661 AA.
AC   Q62167; O09060; O09143;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DEAD-box protein 3 (DEAD-box RNA helicase DEAD3) (mDEAD3) (Embryonic
DE   RNA helicase) (D1PAS1 related sequence 2).
GN   DDX3 OR DEAD3 OR ERH OR D1PAS1-RS2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6, and DBA;
RX   MEDLINE=97104282; PubMed=8948440;
RA   Sowden J.C., Putt W., Morrison K., Beddington R., Edwards Y.;
RT   "The embryonic RNA helicase gene (ERH): a new member of the DEAD box
RT   family of RNA helicases.";
RL   Biochem. J. 308:839-846(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythroleukemia;
RX   MEDLINE=94192995; PubMed=8144024;
RA   Gee S.L., Conboy J.G.;
RT   "Mouse erythroid cells express multiple putative RNA helicase genes
RT   exhibiting high sequence conservation from yeast to mammals.";
RL   Gene 140:171-177(1994).
RN   [3]
RP   SEQUENCE OF 1-9, AND ACETYLATION.
RX   MEDLINE=20318637; PubMed=10859333;
RA   Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,
RA   Lopez de Castro J.A.;
RT   "An N-acetylated natural ligand of human histocompatibility leukocyte
RT   antigen (HLA)-B39. Classical major histocompatibility complex class I
RT   proteins bind peptides with a blocked NH(2) terminus in vivo.";
RL   J. Exp. Med. 191:2083-2092(2000).
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. IT MAY PLAY A ROLE
CC       IN TRANSLATIONAL ACTIVATION OF MRNA IN THE OOCYTE AND EARLY
CC       EMBRYO.
CC   -!- TISSUE SPECIFICITY: DEVELOPMENTALLY REGULATED.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN OOCYTES. UBIQUITOUSLY FOUND IN 9
CC       DAYS POST-CONCEPTION EMBRYO, AT LATER STAGES IT IS RESTRICTED TO
CC       BRAIN AND KIDNEY.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z38117; CAA86261.1; -.
DR   EMBL; L25126; AAA53630.1; -.
DR   HSSP; Q58083; 1HV8.
DR   MGD; MGI:103064; Ddx3.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; DNA-binding; Acetylation.
FT   INIT_MET       0      0
FT   NP_BIND      223    230       ATP (POTENTIAL).
FT   SITE         346    349       DEAD BOX.
FT   DOMAIN       581    661       GLY/SER-RICH.
FT   DOMAIN       608    615       POLY-SER.
FT   DOMAIN       623    629       POLY-GLY.
FT   DOMAIN       632    640       POLY-GLY.
FT   MOD_RES        1      1       ACETYLATION.
SQ   SEQUENCE   661 AA;  72970 MW;  A1E1FAAB5D19F57B CRC64;
```

```
Alignment Scores:
Pred. No.:              3.84e-79      Length:         661
Score:                  1229.00       Matches:        297
Percent Similarity:     57.59%        Conservative:   90
Best Local Similarity:  44.20%        Mismatches:     183
Query Match:            31.08%        Indels:         102
DB:                     1             Gaps:           22

US-09-714-865-15 (1-2172) x DDX3_MOUSE (1-661)

Qy    214 GAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGT--- 270
Db     19 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 36

Qy    271 ---------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
Db     37 TyrIleProProHisLeuArgAsnArgGlu---AlaThrLysGlyPheTyrAspLysAsp 55

Qy    316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGA 375
Db     56 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSer--------------- 70

Qy    376 GGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATAC 435
Db     71 ---PheGlySerArgGlyAspSerArg------GlyLysSerSerPhePheGlyAspArg 87

Qy    436 AGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGCCGTGGAGGATTT---GGTCTA 486
Db     88 GlySerGlySerArgGlyArgPheAspAspArgGly---ArgGlyAspTyrAspGlyIle 106

Qy    487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 546
Db    107 Gly--------------------------------------------------------- 107

Qy    547 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 606
Db    108 ------------------------------GlyArgGlyAsp------------ArgSer 113

Qy    607 GGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGC 663
Db    114 GlyPheGlyLysPheGluArgGlyGly------------AsnSerArgTrpCysAspLys 129

Qy    664 TCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGA 723
Db    130 SerAspGluAspAspTrpSerLys------------------------------------ 137

Qy    724 CCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT---GAGGACTCCATCTTTGCA 780
Db    138 ------------------ProLeuProProSerGluArgLeuGluGlnGluLeuPheSer 151

Qy    781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
Db    152 GlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIleProValGluAlaThrGly 171

Qy    841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
Db    172 AsnAsnCysProProHisIleGluSerPheSerAspValGluMetGlyGluIleIleMet 191

Qy    901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
Db    192 GlyAsnIleGluLeuThrArgTyrThrArgProThrProValGlnLysHisAlaIlePro 211

Qy    961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
Db    212 IleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 231

Qy   1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA------------ATAACT 1068
Db    232 AlaPheLeuLeuProIleLeuSerGlnIleTyrAlaAspGlyProGlyGluAlaLeuArg 251

Qy   1069 GCCAGTCGT---------------TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTA 1113
Db    252 AlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyrProIleSerLeuValLeu 271

Qy   1114 GCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGG 1173
Db    272 AlaProThrArgGluLeuAlaValGlnIleTyrGluGluAlaArgLysPheSerTyrArg 291

Qy   1174 ACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAA 1233
Db    292 SerArgValArgProCysValValTyrGlyGlyAlaGluIleGlyGlnGlnIleArgAsp 311

Qy   1234 ATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGC 1293
Db    312 LeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArgLeuValAspMetMetGlu 331

Qy   1294 AAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATG 1353
Db    332 ArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeuAspGluAlaAspArgMet 351

Qy   1354 TTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTCTTGCCCAGGAATGCCATCA 1413
Db    352 LeuAspMetGlyPheGluProGlnIleArgArgIleValGluGlnAspThrMetProPro 371

Qy   1414 AAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTG 1473
Db    372 LysGlyValArgHisThrMetMetPheSerAlaThrPheProLysGluIleGlnMetLeu 391

Qy   1474 GCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCA 1533
Db    392 AlaArgAspPheLeu---AspGluTyrIlePheLeuAlaValGlyArgValGlySerThr 410

Qy   1534 TGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAGAGAAAAGCTC 1593
Db    411 SerGluAsnIleThrGlnLysValValTrpValGluGluIleAspLysArgSerPheLeu 430

Qy   1594 GTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATGGTCTTTGTTGAAACTAAG 1650
Db    431 LeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeuValPheValGluThrLys 450

Qy   1651 AAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATC 1710
Db    451 LysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGlyTyrAlaCysThrSerIle 470

Qy   1711 CATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAG 1770
Db    471 HisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHisGlnPheArgSerGlyLys 490

Qy   1771 TGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAA 1830
Db    491 SerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeuAspIleSerAsnValLys 510

Qy   1831 CATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGT 1890
Db    511 HisValIleAsnPheAspLeuProSerAspIleGluGluTyrValHisArgIleGlyArg 530

Qy   1891 ACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAAC 1950
Db    531 ThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePheAsn---GluArgAsnIle 549

Qy   1951 CATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGG 2010
Db    550 AsnIleThrLysAspLeuLeuAspLeuLeuValGluAlaLysGlnGluValProSerTrp 569

Qy   2011 TTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAAC 2070
Db    570 LeuGluAsnMetAlaPheGluHisHis---------TyrLysGlySerSerArgGlyArg 586

Qy   2071 GTG-----------------------TTTGCATCAGTTGATACCAGAAAGGGCAAGAGC 2106
Db    587 SerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAspTyrArgGln---SerSer 605

Qy   2107 ACTTTGAACACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
Db    606 GlyAlaSerSerSerSerPheSerSerSerArgAla 617
```

```
RESULT 6
AN3_XENLA
ID   AN3_XENLA      STANDARD;      PRT;   697 AA.
AC   P24346;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ATP-dependent RNA helicase An3.
GN   AN3.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91141586; PubMed=1996140;
RA   Gururajan R., Perry-O'Keefe H., Melton D.A., Weeks D.L.;
RT   "The Xenopus localized messenger RNA An3 may encode an ATP-dependent
RT   RNA helicase.";
RL   Nature 349:717-719(1991).
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X57328; CAA40605.1; -.
DR   PIR; S13654; S13654.
DR   PIR; S29676; S29676.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding.
FT   NP_BIND     265    272       ATP (BY SIMILARITY).
FT   SITE        388    391       DEAD BOX.
FT   DOMAIN      623    697       GLY/SER-RICH.
SQ   SEQUENCE   697 AA;  77302 MW;  F3DD23EB60B2E2EF CRC64;

Alignment Scores:
Pred. No.:               8.1e-79      Length:        697
Score:                   1224.50      Matches:       290
Percent Similarity:      56.85%       Conservative:  96
Best Local Similarity:   42.71%       Mismatches:    198
Query Match:             30.97%       Indels:        95
DB:                      1            Gaps:          18

US-09-714-865-15 (1-2172) x AN3_XENLA (1-697)

Qy    181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
            :::  ||||||     |||   |||:::|||
Db     47 AlaSerArgAsnAspSerAsnTrpAspSerGlyArg------------------------ 58

Qy    241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
                                             |||     |||:::   |||
Db     59 --------------------------------GlyGlyAsnGlyTyrIleAsnGlyMet 67

Qy    301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
             :::|||   |||                 |||   :::       |||
Db     68 GlnAspAspArgAspGly------------ArgMetAsnGlyTyrAspArgGlyGlyTyr 83

Qy    361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
             :::|||     |||  |||  ||||||  |||   |||        ||||||
Db     84 GlySerArgGlyThrGlyArgSerAspArgGlyPheTyr---AspArgGluAsnSerGly 102

Qy    421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
             :::            :::         ||||||  |||   ||||||
Db    103 TrpAsnSerGlyArgAspLysAspAlaTyrSerSerPhe---GlySerArgGlyAspArg 121

Qy    481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
          |||   ||||||
Db    122 GlyLysGlySer------------------------------------------------ 125

Qy    541 CTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
          ||||||      |||                |||:::|||:::      |||
Db    126 LeuPheAsnGluArg---------------GlySerGlySerArgArgThrAspAspArg 140

Qy    601 AGAAGT------------------GGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTA 642
          |||                     |||:::  |||:::::   |||:::
Db    141 ArgGlnAspGlyPheAspGlyMetGlyAsnArgSerAspLysSerGlyPheGlyArgPhe 160

Qy    643 AATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGA 702
          :::                    |||   :::   |||
Db    161 AspArg-----------------GlyAsnSerArgTrp--------------------- 167

Qy    703 GAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT 762
                |||||   :::      :::      :::         |||       |||   |||
Db    168 ------SerAspAspArgAsnAspGluAspAspTrpSerLysProLeuAlaProAsnAsp 185

Qy    763 ------GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGAC 816
            |||      :::|||:::           ||||||||||||:::||||||||
Db    186 ArgValGluGlnGluLeuPheSerGlySerAsnThrGlyIleAsnPheGluLysTyrAsp 205

Qy    817 ACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAA 876
             |||   ||||||   :::|||    :::    ||||||   |||     :::|||    :::
Db    206 AspIleProValGluAlaThrGlySerAsnCysProProHisIleGluSerPheHisAsp 225

Qy    877 GCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACT 936
             :::   :::   :::       ||||||              ||||||:::   |||
Db    226 ValThrMetGlyGluIleIleMetGlyAsnIleGlnLeuThrArgTyrThrArgProThr 245

Qy    937 CCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCT 996
          ||||||||||||::::::||||||||||||:::      ||||||||||||||||||
Db    246 ProValGlnLysHisAlaIleProIleIleIleGluLysArgAspLeuMetAlaCysAla 265

Qy    997 CAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCAT 1056
          ||||||||||||||||||||||||||||||||||||||||||:::   :::
Db    266 GlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrAla 285

Qy   1057 GATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA-------------------- 1095
          ||||||            |||   |||||||||
Db    286 AspGly---ProGlyAspAlaMetLysHisLeuGlnGluAsnGlyArgTyrGlyArgArg 304

Qy   1096 ---------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
                   |||      :::::::::||||||||||||||||||      |||||||||
Db    305 LysGlnPheProLeuSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyr 324

Qy   1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
             ||||||||||||||:::::      :::   ||||||      |||:::|||||||||
Db    325 GluGluAlaArgLysPheAlaTyrArgSerArgValArgProCysValValTyrGlyGly 344

Qy   1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
             :::|||       ||||||    :::      :::||||||::::::|||    ||||||
Db    345 AlaAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThr 364

Qy   1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
          ||||||||||||:::|||::::::      :::      ||||||||||||      ||||||
Db    365 ProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyr 384

Qy   1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
          ||||||||||||||||||||||||||||||||||||||||||      |||::::::::::::
Db    385 LeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArg 404
```

```
Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              ::::::           ||||||   |||       ||||||||:::|||||||||||
Db        405 IleValGluGlnAspThrMetProProLysGlyValArgGlnThrMetMetPheSerAla 424

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              |||||||||:::|||||||||  ||||||   :::||||||           |||:::|||
Db        425 ThrPheProLysGluIleGlnIleLeuAlaArgAspPheLeu---AspGluTyrIlePhe 443

Qy       1507 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1566
              :::|||||||||:::||||||            ::::::  |||  |||:::   |||
Db        444 LeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpVal 463

Qy       1567 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAA 1623
                 :::        ||||||     |||:::::::::|||         |||   |||
Db        464 GluGluMetAspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSer 483

Qy       1624 AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGT 1683
                 |||:::||||||||||||||||||||||   ||||||          ||||||
Db        484 LeuThrLeuValPheValGluThrLysLysGlyAlaAspAlaLeuGluAspPheLeuTyr 503

Qy       1684 CAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAA 1743
                 |||      :::   ||||||||||||||||||   |||||||:::|||||||:::
Db        504 HisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGlu 523

Qy       1744 GCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCC 1803
              ||||||      ||||||    ||||||   |||:::||||||||||||:::|||||||||
Db        524 AlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAla 543

Qy       1804 AGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATT 1863
              ||||||||||||||     ||||||:::|||||||||||||||||||||||||    |||
Db        544 ArgGlyLeuAspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIle 563

Qy       1864 GATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATT 1923
              :::|||||||||||||||||||||||||||||||||   ||||||   |||    |||
Db        564 GluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThr 583

Qy       1924 TCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACA 1983
              ||||||||||:::   |||    :::   ::::::   :::    |||:::    :::|||
Db        584 SerPhePheAsn---GluLysAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuVal 602

Qy       1984 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCT 2043
              :::|||:::|||:::||||||:::|||||||||    :::|||:::        :::
Db        603 GluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaTyrGluGlnHis------ 620

Qy       2044 GGCTTCAGTGGTAGTACAAGAGGAAACGTG--------------------TTTGCATCA 2082
                          |||:::||||||                         |||    :::
Db        621 ---HisLysSerSerSerArgGlyArgSerLysSerArgPheSerGlyGlyPheGlyAla 639

Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
                 |||    |||:::                   :::    |||   ||||||||
Db        640 LysAspTyrArgGlnSerSerGlyAla------GlySerSerPheGlySerSerArg 656

RESULT 7
DDX3_HUMAN
ID   DDX3_HUMAN     STANDARD;      PRT;   661 AA.
AC   O00571; O15536;
DT   15-JUL-1998 (Rel. 36, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 3 (Helicase-like protein 2) (HLP2) (DEAD-box, X
DE   isoform).
GN   DDX3 OR DBX.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver, and Hippocampus;
RA   Chung J., Lee S.-G., Song K.;
RT   "Identification of a human homolog of a putative RNA helicase gene
RT   (mDEAD3) expressed in mouse erythroid cells.";
RL   Korean J. Biochem. 27:193-197(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Owsianka A.M., Patel A.H.;
RT   "DEAD box putative RNA helicase.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98022381; PubMed=9381176;
RA   Lahn B.T., Page D.C.;
RT   "Functional coherence of the human Y chromosome.";
RL   Science 278:675-680(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 1-9, AND ACETYLATION.
RX   MEDLINE=20318637; PubMed=10859333;
RA   Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,
RA   Lopez de Castro J.A.;
RT   "An N-acetylated natural ligand of human histocompatibility leukocyte
RT   antigen (HLA)-B39. Classical major histocompatibility complex class I
RT   proteins bind peptides with a blocked NH(2) terminus in vivo.";
RL   J. Exp. Med. 191:2083-2092(2000).
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. INTERACTS
CC       SPECIFICALLY WITH HEPATITIS C VIRUS CORE PROTEIN RESULTING A
CC       CHANGE IN INTRACELLULAR LOCATION.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U50553; AAB95637.1; -.
DR   EMBL; AF061337; AAC34298.1; -.
DR   EMBL; AF000983; AAC51830.1; -.
DR   EMBL; AF000982; AAC51829.1; -.
DR   EMBL; BC011819; AAH11819.1; -.
DR   HSSP; Q58083; 1HV8.
DR   Genew; HGNC:2745; DDX3.
DR   MIM; 300160; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; DNA-binding; Acetylation.
FT   INIT_MET      0      0
FT   NP_BIND     223    230       ATP (POTENTIAL).
FT   SITE        346    349       DEAD BOX.
FT   DOMAIN      581    661       GLY/SER-RICH.
FT   MOD_RES       1      1       ACETYLATION.
FT   CONFLICT     49     49       K -> R (IN REF. 3).
SQ   SEQUENCE   661 AA;  73112 MW;  F0F03DD8FBC00A65 CRC64;

Alignment Scores:
Pred. No.:              3.74e-78       Length:         661
Score:                  1215.00        Matches:        293
Percent Similarity:     56.64%         Conservative:   91
Best Local Similarity:  43.22%         Mismatches:     180
Query Match:            30.73%         Indels:         114
DB:                     1              Gaps:           20

US-09-714-865-15 (1-2172) x DDX3_HUMAN (1-661)
```

```
Qy        214 GAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGT--- 270
              :::   |||      |||||   |||      |||||        |||
Db         19 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 36

Qy        271 --------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
                                |||||      :::        |||   |||    |||
Db         37 TyrIleProProHisLeuArgAsnArgGlu---AlaThrLysGlyPheTyrAspLysAsp 55

Qy        316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------------------GAA 354
              ||||||||:::                 |||
Db         56 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 75

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
              |||:::      ::::::      |||||   |||||
Db         76 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer----------------- 89

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGC 468
                                       |||   ||||||   |||       ||||||
Db         90 --------------------------GlySerArgGlyArgPheAspAspArgGlyArg 100

Qy        469 CGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG 528
                         |||:::||||||
Db        101 SerAspTyrAspGlyIleGlySer------------------------------------ 108

Qy        529 CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGAT 588
                                                                   ||||||
Db        109 ---------------------------------------------------ArgGlyAsp 111

Qy        589 ACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAAT 645
                         ||||||||  ·|||      |||||||||||            |||
Db        112 ------------ArgSerGlyPheGlyLysPheGluArgGlyGly-----------Asn 123

Qy        646 GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
                          |||   ::::::   |||
Db        124 SerArgTrpCysAspLysSerAspGluAspAspTrpSerLys----------------- 137

Qy        706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--- 762
                                                  |||   ||||||   |||
Db        138 -----------------------------------ProLeuProProSerGluArgLeu 145

Qy        763 GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATT 822
              |||      :::|||:::        ||||||||||||||:::|||||||||    |||
Db        146 GluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIle 165

Qy        823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                 ||||||   :::|||::::::   ||||||   |||    :::|||    :::
Db        166 ProValGluAlaThrGlyAsnAsnCysProProHisIleGluSerPheSerAspValGlu 185

Qy        883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
              :::   :::   :::                     |||||||:::   |||||||||
Db        186 MetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgProThrProVal 205

Qy        943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
              ||||||::::::||||||||||||         ||||||||||||||||||||||||
Db        206 GlnLysHisAlaIleProIleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThr 225

Qy       1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA 1062
              ||||||||||||||||||||||||||||||||||||:::     :::      ||||||
Db        226 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 245

Qy       1063 ------------ATAACTGCCAGTCGT---------------TTTAAAGAGTTGCAGGAA 1095
                          :::   |||   :::                :::       |||
Db        246 ProGlyGluAlaLeuArgAlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyr 265

Qy       1096 CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCC 1155
              |||      :::::::::||||||||||||||||||      |||||||||   ||||||
Db        266 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluGluAla 285

Qy       1156 AGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTG 1215
              ||||||||||||:::   :::   ||||||     |||:::||||||||||    :::
Db        286 ArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGlyAlaAspIle 305

Qy       1216 GGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGA 1275
              |||      ||||||   :::     :::||||||::::::|||   ||||||||||||||
Db        306 GlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArg 325

Qy       1276 CTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTG 1335
              |||:::|||::::::   :::              |||||||||||||||
Db        326 LeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeu 345

Qy       1336 GATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT 1395
              |||||||||||||||||||||||||||||||     |||::::::::::::::::::
Db        346 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArgIleValGlu 365

Qy       1396 TGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCA 1455
                       ||||||   |||         |||   |||:::||||||||||||||||||
Db        366 GlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAlaThrPhePro 385

Qy       1456 GAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT 1515
              :::||||||||||   ||||||   :::||||||               |||:::|||:::||||||
Db        386 LysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePheLeuAlaVal 404

Qy       1516 GGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTC 1575
              |||:::||||||      ::::::   |||   |||:::     |||      :::
Db        405 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpValGluGluSer 424

Qy       1576 TCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATG 1632
                    ||||||      |||:::::::::|||         |||   |||    |||:::
Db        425 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeu 444

Qy       1633 GTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAA 1692
              ||||||||||||||||||      ||||||               ||||||    |||
Db        445 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGly 464

Qy       1693 ATATCAACTACAAGTATCCATGGTGATCGGAACAGAGAGAGCGGGAGCAAGCTCTTGGA 1752
                 :::   ||||||||||||||||||   ||||||:::||||||:::||||||
Db        465 TyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHis 484

Qy       1753 GATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTG 1812
                 ||||||   ||||||      |||:::||||||||||||:::||||||||||||||||
Db        485 GlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 504

Qy       1813 GATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATAT 1872
              ||||||   ||||||:::||||||||||||||||||||||||||||||   |||:::||||||
Db        505 AspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIleGluGluTyr 524

Qy       1873 GTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTT 1932
              ||||||||||||||||||||||||||||    ||||||   |||   |||   |||||||||
Db        525 ValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 544

Qy       1933 GATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAA 1992
              :::   |||   :::   ::::::   :::   |||:::   :::|||   :::|||:::
Db        545 Asn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuValGluAlaLys 563

Qy       1993 CAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGT 2052
              |||:::||||||:::|||||||||      :::|||:::          :::
Db        564 GlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis---------TyrLys 580

Qy       2053 GGTAGTACAAGAGGAAACGTG-----------------------TTTGCATCAGTTGAT 2088
              ||||||:::||||||                             |||   :::   |||
Db        581 GlySerSerArgGlyArgSerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAsp 600

Qy       2089 ACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                 |||:::      |||      :::::::::   ||||||||||||||||||
Db        601 TyrArgGln---SerSerGlyAlaSerSerSerSerPheSerSerSerArgAla 617
```

RESULT 8
PL10_MOUSE
ID   PL10_MOUSE   STANDARD;   PRT;   660 AA.
AC   P16381;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ATP-dependent RNA helicase PL10.
GN   D1PAS1 OR PL10.

```
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=89249320; PubMed=2720782;
RA   Leroy P., Alzari P., Sassoon D., Wolgemuth D., Fellous M.;
RT   "The protein encoded by a murine male germ cell-specific transcript
RT   is a putative ATP-dependent RNA helicase.";
RL   Cell 57:549-559(1989).
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. POSSIBLE ROLE
CC       IN A KEY STEP OF THE SPERMATOGENIC PROCESS.
CC   -!- TISSUE SPECIFICITY: TESTIS.
CC   -!- DEVELOPMENTAL STAGE: HIGH LEVELS OF PL10 DURING THE MEIOTIC AND
CC       HAPLOID STAGES OF SPERMATOGENESIS.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J04847; AAA39942.1; -.
DR   PIR; A32378; A32378.
DR   HSSP; Q58083; 1HV8.
DR   MGD; MGI:91842; D1Pas1.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; DNA-binding; Spermatogenesis.
FT   NP_BIND     223    230       ATP (POTENTIAL).
FT   SITE        346    349       DEAD BOX.
FT   DNA_BIND    495    514       POTENTIAL.
FT   DOMAIN      581    660       GLY/SER-RICH.
SQ   SEQUENCE   660 AA;  73140 MW;  50AD4E6A131AE603 CRC64;

Alignment Scores:
Pred. No.:              6.97e-77      Length:         660
Score:                  1197.00       Matches:        286
Percent Similarity:     57.18%        Conservative:   104
Best Local Similarity:  41.94%        Mismatches:     186
Query Match:            30.27%        Indels:         106
DB:                     1             Gaps:           19

US-09-714-865-15 (1-2172) x PL10_MOUSE (1-660)

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              :::|||   :::       :::||||||:::                   :::|||:::
Db         17 AlaGlyLeuAspLeuThrSerArgAspSer---------GlnSerGlyGlySerThrAla 33

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTCAAACAGCAGG 300
              :::    |||   :::   :::                |||||      :::   :::
Db         34 SerLysGlyArgTyr---IleProProHisLeuArgAsnArgGluAlaAlaLysAla--- 51

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAG---------------TCTAGTAAT 345
              |||   |||.  |||   |||    :::   ::::::                   |||
Db         52 PheTyrAspLysAspGlySerArgTrpSerLysAspLysAspAlaTyrSerSerPheGly 71

Qy        346 GACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGAT 405
                       |||       ::::::      ||||||    |||||||||
Db         72 SerArgSerAspThrArgAlaLysSerSerPhePheSerAspArgGlyGly--------- 88

Qy        406 GGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGA--- 462
                       |||   ::: . |||     :::         ||||||         :::
Db         89 ---------SerGlySerArgGlyArgPheAspGluArgGlyArgSerAspTyrGluSer 105

Qy        463 ---GGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAA 519
                 |||   |||||||||
Db        106 ValGlySerArgGlyGly------------------------------------------ 111

Qy        520 TGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGT 579

Db        111 ------------------------------------------------------------ 111

Qy        580 AATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTACAAA 636
                                   |||||||||   |||     |||||||||
Db        112 ---------------------ArgSerGlyPheGlyLysPheGluArgGly--------- 121

Qy        637 GGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAA 696
                                             |||   :::   |||      :::|||:::
Db        122 ------------------------------GlyAsnSerArgTrpCysAspLysAlaAsp 131

Qy        697 GGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCT 756
                 :::   |||                                 |||   ||||||
Db        132 GluAspAspTrpSerLys--------------------------ProLeuProProSer 142

Qy        757 GAGGAT---GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATAC 813
              |||       |||      :::|||:::          ||||||||||||||:::||||||
Db        143 GluArgLeuGluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyr 162

Qy        814 GACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAA 873
              |||  . |||   ||||||   :::|||::::::   ||||||   |||   :::|||
Db        163 AspAspIleProValGluAlaThrGlyAsnAsnCysProProHisIleGluSerPheSer 182

Qy        874 GAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTT 933
              :::      :::   :::   :::         ||||||            ||||||:::
Db        183 AspValGluMetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgPro 202

Qy        934 ACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGT 993
              |||||||||||||||||:::::::||||||||||||      ||||||||||||||||||
Db        203 ThrProValGlnLysHisAlaIleProIleIleLysGluLysArgAspLeuMetAlaCys 222

Qy        994 GCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATG 1053
              ||||||||||||||||||||||||||||||||||||||||||||||||:::   :::
Db        223 AlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyr 242

Qy       1054 CATGATGGA------------ATAACTGCCAGTCGT---------------TTTAAAGAG 1086
                 ||||||          :::   |||   :::                  :::
Db        243 ThrAspGlyProGlyGluAlaLeuArgAlaMetLysGluAsnGlyLysTyrGlyArgArg 262

Qy       1087 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
                 |||   |||         ::::::::||||||||||||||||||      |||||||||
Db        263 LysGlnTyrProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyr 282

Qy       1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
              ||||||||||||||||||:::         :::         |||:::|||||||||
Db        283 GluGluAlaArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGly 302

Qy       1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
              :::|||         ||||||   :::   :::|||||||::::::|||      ||||||
Db        303 AlaAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThr 322

Qy       1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
              ||||||||||||:::|||::::::   :::   |||||||||||||         ||||||
Db        323 ProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyr 342

Qy       1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
              ||||||||||||||||||||||||||||||||||||||||||||   |||::::::::::::
Db        343 LeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArg 362

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              ::::::            ||||||   |||         |||   |||:::||||||||||||
Db        363 IleValGluGlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAla 382

Qy       1447 ACTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTAAAGTCAAATTATCTGTTT 1506
```

```
                ||||||||:::||||||||  ||||||   :::||||||        |||:::|||
Db          383 ThrPheProLysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePhe 401

Qy         1507 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1566
                :::|||||||||:::||||||            ::::::   |||   |||:::   |||
Db          402 LeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpVal 421

Qy         1567 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAA 1623
                   :::      ||||||      |||::::::::|||          |||   |||
Db          422 GluGluAlaAspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSer 441

Qy         1624 AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGT 1683
                      :::||||||||||||||||||    ||||||              ||||||
Db          442 LeuIleLeuValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyr 461

Qy         1684 CAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAA 1743
                   |||      :::   ||||||||||||||||||   ||||||:::||||||:::
Db          462 HisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGlu 481

Qy         1744 GCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCC 1803
                ||||||      ||||||   ||||||   |||:::|||||||||||:::|||||||||
Db          482 AlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAla 501

Qy         1804 AGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATT 1863
                ||||||||||||||   ||||||:::||||||||||||||||||||||||      |||
Db          502 ArgGlyLeuAspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIle 521

Qy         1864 GATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATT 1923
                :::||||||||||||||||||||||||||||||   ||||||   |||   |||
Db          522 GluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThr 541

Qy         1924 TCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACA 1983
                |||||||||:::   |||   :::   ::::::   :::   |||:::   :::|||
Db          542 SerPhePheAsn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuVal 560

Qy         1984 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCT 2043
                :::|||:::|||:::||||||:::|||||||||   :::||||||      :::
Db          561 GluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaPheGluHisHis------ 578

Qy         2044 GGCTTCAGTGGTAGTACAAGAGGAAACGTG-----------------------TTTGCATCA 2082
                   :::   |||   :::||||||                        |||   :::
Db          579 ---TyrLysGlyGlySerArgGlyArgSerLysSerArgPheSerGlyGlyPheGlyAla 597

Qy         2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                   |||   |||:::      |||      :::::::::   |||||||||   ||||||
Db          598 ArgAspTyrArgGln---SerSerGlyAlaSerSerSerSerPheSerSerGlyArgAla 616

Qy         2143 CCCAAT 2148
                   |||
Db          617 SerAsn 618

RESULT 9
DDXY_HUMAN
ID   DDXY_HUMAN     STANDARD;      PRT;   660 AA.
AC   O15523;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 3, Y-chromosomal.
GN   DBY.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98022381; PubMed=9381176;
RA   Lahn B.T., Page D.C.;
RT   "Functional coherence of the human Y chromosome.";
RL   Science 278:675-680(1997).
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF000985; AAC51832.1; -.
DR   EMBL; AF000984; AAC51831.1; -.
DR   HSSP; Q58083; 1HV8.
DR   Genew; HGNC:2699; DBY.
DR   MIM; 400010; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; DNA-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
FT   SITE        345    348       DEAD BOX.
SQ   SEQUENCE   660 AA;  73094 MW;  5A67DC082AEE9CFF CRC64;

Alignment Scores:
Pred. No.:               7.36e-76      Length:        660
Score:                   1182.50       Matches:       278
Percent Similarity:      58.06%        Conservative:  93
Best Local Similarity:   43.51%        Mismatches:    167
Query Match:             29.91%        Indels:        101
DB:                      1             Gaps:          18

US-09-714-865-15 (1-2172) x DDXY_HUMAN (1-660)

Qy          301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------- 351
                |||   |||   ||||||||||||:::                     |||
Db           52 PheHisAspLysAspSerSerGlyTrpSerCysSerLysAspLysAspAlaTyrSerSer 71

Qy          352 ---------GAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
                         |||:::         :::         ||||||:::||||||
Db           72 PheGlySerArgAspSerArgGlyLysProGlyTyrPheSerGluArgGlySer------ 89

Qy          403 GATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC--- 459
                                                     |||   ||||||   |||
Db           90 ---------------------------------------GlySerArgGlyArgPheAsp 96

Qy          460 ---CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGAC 516
                   ||||||                              :::|||   |||
Db           97 AspArgGly-----------------------------ArgSerAspTyrAsp------ 104

Qy          517 GAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACA 576
                                                                     |||
Db          105 ------------------------------------------------------GlyIle 106

Qy          577 GGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTAC 633
                ||||||            :::   |||   |||   |||      ||||||   |||:::
Db          107 GlyAsn------------ArgGluArgProGlyPheGlyArgPheGluArgSerGlyHis 122

Qy          634 AAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCA 693
                         ::::::            |||   ::::::   |||
Db          123 SerArgTrpCysAspLys------------SerValGluAspAspTrpSerLys------ 136

Qy          694 GAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCA 753
                                                                |||   ||||||
Db          137 ------------------------------------------------ProLeuProPro 140

Qy          754 CCTGAGGAT---GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAA 810
                   |||      |||      :::|||:::            ||||||||||||:::|||
Db          141 SerGluArgLeuGluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLys 160

Qy          811 TACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTT 870
```

```
              ||||||  |||  ||||||  :::|||  :::  ||||||  |||       |||
Db      161 TyrAspAspIleProValGluAlaThrGlySerAsnCysProProHisIleGluAsnPhe 180

Qy      871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
              :::  ::::::  :::  :::           ||||||          ||||||:::
Db      181 SerAspIleAspMetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArg 200

Qy      931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
              ||||||||||||||::::::|||||||||||              |||||||||:::|||
Db      201 ProThrProValGlnLysHisAlaIleProIleIleLysGlyLysArgAspLeuValAla 220

Qy      991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||:::      :::
Db      221 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIle 240

Qy     1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA--------------- 1095
                    ||||||                     |||    ::::::|||
Db      241 TyrThrAspGly---ProGlyGluAlaLeuLysAlaValLysGluAsnGlyArgTyrGly 259

Qy     1096 ---------------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
                           |||      :::::::::||||||||||||||||||       |||
Db      260 ArgArgLysGlnTyrProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGln 279

Qy     1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
              ||||||   ||||||||||||||||:::   :::   ||||||       |||:::|||
Db      280 IleTyrGluGluAlaArgLysPheSerTyrArgSerArgValArgProCysValValTyr 299

Qy     1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
              ||||||      :::|||       ||||||   :::      :::||||||::::::|||
Db      300 GlyGlyAlaAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuVal 319

Qy     1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
              ||||||||||||||||||:::|||::::::   :::   ||||||||||||
Db      320 AlaThrProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCys 339

Qy     1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||::::::
Db      340 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIle 359

Qy     1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
              :::::::::::               ||||||  |||     |||  |||:::||||||
Db      360 ArgArgIleValGluGlnAspThrMetProProLysGlyValArgHisThrMetMetPhe 379

Qy     1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
              ||||||||||||||:::|||||||||   ||||||    :::||||||       |||
Db      380 SerAlaThrPheProLysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyr 398

Qy     1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
              :::|||:::|||||||||:::||||||            ::::::  |||   |||:::
Db      399 IlePheLeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValVal 418

Qy     1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG--- 1617
                 |||            ||||||    |||:::::||||||          |||
Db      419 TrpValGluAspLeuAspLysArgSerPheLeuLeuAspIleLeuGlyAlaThrGlySer 438

Qy     1618 GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTT 1677
              |||       |||:::||||||||||||||||||||    ||||||          |||
Db      439 AspSerLeuThrLeuValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPhe 458

Qy     1678 CTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGG 1737
              |||       |||    :::   |||||||||||||||||||||    ||||||:::|||
Db      459 LeuTyrHisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArg 478

Qy     1738 GAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTA 1797
              |||:::||||||      ||||||  ||||||   |||:::||||||||||||:::|||
Db      479 GluGluAlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaVal 498

Qy     1798 GCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCT 1857
              |||||||||||||||||||     ||||||:::|||||||||||||||||||||||||||
Db      499 AlaAlaArgGlyLeuAspIleSerAsnValArgHisValIleAsnPheAspLeuProSer 518

Qy     1858 ACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGA 1917
                 |||:::||||||||||||||||||||||||||||||||    ||||||   |||
Db      519 AspIleGluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeu 538

Qy     1918 GCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTA 1977
              |||   |||||||||:::   |||   :::   ::::::   :::   |||:::   :::
Db      539 AlaThrSerPhePheAsn---GluLysAsnMetAsnIleThrLysAspLeuLeuAspLeu 557

Qy     1978 TTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATAC 2037
              |||   :::|||:::|||:::||||||:::|||||||||     :::|||:::      :::
Db      558 LeuValGluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis 577

Qy     2038 ATTCCTGGCTTCAGT---GGTAGTACAAGAGGAAACGTGTTTGCA-----------TCA 2082
                       |||    |||    |||     ::::::   |||    |||:::       :::
Db      578 TyrLysGlyGlySerArgGlyArgSerLysSerAsnArgPheSerGlyGlyPheGlyAla 597

Qy     2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
                 |||    |||:::         :::    :::::::::||||||    :::||||||
Db      598 ArgAspTyrArgGlnSerSerGlySer---SerSerSerGlyPheGlyAlaSerArg 615
```

```
RESULT 10
DED1_YEAST
ID   DED1_YEAST     STANDARD;      PRT;   604 AA.
AC   P06634;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable ATP-dependent RNA helicase DED1.
GN   DED1 OR SPP81 OR YOR204W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A364A X H79-20.3;
RX   MEDLINE=91141585; PubMed=1996139;
RA   Jamieson D.J., Rahe B., Pringle J., Beggs J.D.;
RT   "A suppressor of a yeast splicing mutation (prp8-1) encodes a
RT   putative ATP-dependent RNA helicase.";
RL   Nature 349:715-717(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Hughes B., Pohl T.M.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-112 FROM N.A.
RX   MEDLINE=86093663; PubMed=3001645;
RA   Struhl K.;
RT   "Nucleotide sequence and transcriptional mapping of the yeast
RT   pet56-his3-ded1 gene region.";
RL   Nucleic Acids Res. 13:8587-8601(1985).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=97289744; PubMed=9144215;
RA   de la Cruz J., Iost I., Kressler D., Linder P.;
RT   "The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent
RT   translation in Saccharomyces cerevisiae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:5201-5206(1997).
CC   -!- FUNCTION: POSITIVELY INVOLVED IN THE INITIATION OF PROTEIN
CC       SYNTHESIS. PROBABLE ATP-BINDING RNA HELICASE. ESSENTIAL FOR CELL
CC       VIABILITY. SUPPRESSOR OF A PRP8 MUTATION. MAY PLAY A ROLE IN MRNA
CC       SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X57278; CAA40546.1; -.
```

```
DR   EMBL; Z75110; CAA99419.1; -.
DR   EMBL; X03245; CAA27004.1; -.
DR   PIR; S13653; S13653.
DR   PIR; S07683; S07683.
DR   HSSP; Q58083; 1HV8.
DR   SGD; S0005730; DED1.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding.
FT   NP_BIND      186    193       ATP (BY SIMILARITY).
FT   SITE         306    309       DEAD BOX.
FT   CONFLICT      37     37       S -> M (IN REF. 2).
SQ   SEQUENCE   604 AA;  65553 MW;  B6722D94C03BFA4B CRC64;

Alignment Scores:
Pred. No.:                6.45e-72     Length:        604
Score:                    1126.50      Matches:       261
Percent Similarity:       56.66%       Conservative:  92
Best Local Similarity:    41.89%       Mismatches:    187
Query Match:              28.49%       Indels:        83
DB:                       1            Gaps:          16

US-09-714-865-15 (1-2172) x DED1_YEAST (1-604)

Qy    337 TCTAGTAATGACTGCGAAGATAAT---------CCAACACGGAACAGAGGGTTTTCCAAG 387
Db     12 SerIleAsnAspAsnAsnGluAsnGlyTyrValProProHisLeuArgGlyLysProArg 31

Qy    388 AGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGT--- 444
Db     32 SerAlaArgAsnAsnSerSerAsnTyrAsnAsnAsnAsnGlyGlyTyrAsnGlyGlyArg 51

Qy    445 GGAAGAGGTAGTTTC------CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAAT 498
Db     52 GlyGlyGlySerPhePheSerAsnAsnArgArgGlyGlyTyrGly--------------- 66

Qy    499 AATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTGGTTCTAGAAGA 558
Db     67 ---------------------------------AsnGlyGlyPhePheGly--------- 72

Qy    559 CCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT 618
Db     73 ------------GlyAsnAsnGlyGlySerArgSerAsnGlyArgSerGly--------- 85

Qy    619 GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT 678
Db     86 ---------------------------------------------------------GlyArg 87

Qy    679 TGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTAC 738
Db     88 Trp---------IleAspGly--------------------------------LysHis 93

Qy    739 ATACCCCCTCCTCCACCTGAGGAT---GAGGACTCCATCTTTGCA--------------- 780
Db     94 ValProAlaProArgAsnGluLysAlaGluIleAlaIlePheGlyValProGluAspPro 113

Qy    781 CATTATCAG---ACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
Db    114 AsnPheGlnSerSerGlyIleAsnPheAspAsnTyrAspAspIleProValAspAlaSer 133

Qy    838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
Db    134 GlyLysAspValProGluProIleThrGluPheThrSerProProLeuAspGlyLeuLeu 153

Qy    898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
Db    154 LeuGluAsnIleLysLeuAlaArgPheThrLysProThrProValGlnLysTyrSerVal 173

Qy    958 CCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACT 1017
Db    174 ProIleValAlaAsnGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThr 193

Qy   1018 GCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACT--------- 1068
Db    194 GlyGlyPheLeuPheProValLeuSerGluSerPheLysThrGlyProSerProGlnPro 213

Qy   1069 ---------GCCAGTCGTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCA 1119
Db    214 GluSerGlnGlySerPheTyrGlnArgLysAlaTyrProThrAlaValIleMetAlaPro 233

Qy   1120 ACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGT 1179
Db    234 ThrArgGluLeuAlaThrGlnIlePheAspGluAlaLysLysPheThrTyrArgSerTrp 253

Qy   1180 GTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTA 1239
Db    254 ValLysAlaCysValValTyrGlyGlySerProIleGlyAsnGlnLeuArgGluIleGlu 273

Qy   1240 CAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAA 1299
Db    274 ArgGlyCysAspLeuLeuValAlaThrProGlyArgLeuAsnAspLeuLeuGluArgGly 293

Qy   1300 AAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGAT 1359
Db    294 LysIleSerLeuAlaAsnValLysTyrLeuValLeuAspGluAlaAspArgMetLeuAsp 313

Qy   1360 ATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAA 1419
Db    314 MetGlyPheGluProGlnIleArgHisIleValGluAspCysAspMetThrProValGly 333

Qy   1420 CAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCA 1479
Db    334 GluArgGlnThrLeuMetPheSerAlaThrPheProAlaAspIleGlnHisLeuAlaArg 353

Qy   1480 GAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGA 1539
Db    354 AspPheLeu---SerAspTyrIlePheLeuSerValGlyArgValGlySerThrSerGlu 372

Qy   1540 GATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAA 1599
Db    373 AsnIleThrGlnLysValLeuTyrValGluAsnGlnAspLysLysSerAlaLeuLeuAsp 392

Qy   1600 ATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCA 1659
Db    393 LeuLeuSerAlaSerThrAspGlyLeuThrLeuIlePheValGluThrLysArgMetAla 412

Qy   1660 GATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGAT 1719
Db    413 AspGlnLeuThrAspPheLeuIleMetGlnAsnPheArgAlaThrAlaIleHisGlyAsp 432

Qy   1720 CGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTT 1779
Db    433 ArgThrGlnSerGluArgGluArgAlaLeuAlaAlaPheArgSerGlyAlaAlaThrLeu 452

Qy   1780 CTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATC 1839
Db    453 LeuValAlaThrAlaValAlaAlaArgGlyLeuAspIleProAsnValThrHisValIle 472

Qy   1840 AATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
Db    473 AsnTyrAspLeuProSerAspValAspAspTyrValHisArgIleGlyArgThrGlyArg 492

Qy   1900 TGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCA 1959
Db    493 AlaGlyAsnThrGlyLeuAlaThrAlaPhePheAsnSerGluAsnSerAsn---IleVal 511

Qy   1960 CAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAA 2019
Db    512 LysGlyLeuHisGluIleLeuThrGluAlaAsnGlnGluValProSerPheLeuLysAsp 531
```

```
Qy       2020 ATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCA 2079
                 |||         ||||||        ::::::|||         |||
Db        532 AlaMetMetSerAla------ProGlySerArgSerAsnSerArgArgGlyGlyPheGly 549

Qy       2080 TCAGTTGATACCAGAAAGGGCAAGAGCACT------TTGAACACAGCTGGGTTTTCTTCT 2133
                 :::    |||      :::          :::       |||:::    |||
Db        550 ArgAsnAsnAsnArgAspTyrArgLysAlaGlyGlyAlaSerAlaGlyGlyTrpGlySer 569

Qy       2134 TCACGAGCT 2142
              ||||||:::
Db        570 SerArgSer 572
```

RESULT 11

```
DED1_SCHPO
ID   DED1_SCHPO     STANDARD;      PRT;   636 AA.
AC   O13370; O59857;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ATP-dependent RNA helicase ded1.
GN   DED1 OR SUM3 OR DEP1 OR MOC2 OR SPCC1795.11.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, AND ALTERNATIVE FORMS.
RX   MEDLINE=21659741; PubMed=11711540;
RA   Liu H.-Y., Nefsky B.S., Walworth N.C.;
RT   "The ded1 DEAD box helicase interacts with Chk1 and Cdc2.";
RL   J. Biol. Chem. 277:2637-2643(2002).
RN   [2]
RP   SEQUENCE FROM N.A., AND FUNCTION.
RX   MEDLINE=20191909; PubMed=10725227;
RA   Grallert B., Kearsey S.E., Lenhard M., Carlson C.R., Nurse P.,
RA   Boye E., Labib K.;
RT   "A fission yeast general translation factor reveals links between
RT   protein synthesis and cell cycle controls.";
RL   J. Cell Sci. 113:1447-1458(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99326140; PubMed=10395922;
RA   Kawamukai M.;
RT   "Isolation of a novel gene, moc2, encoding a putative RNA helicase as
RT   a suppressor of sterile strains in Schizosaccharomyces pombe.";
RL   Biochim. Biophys. Acta 1446:93-101(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99054689; PubMed=9832516;
RA   Forbes K.C., Humphrey T., Enoch T.;
RT   "Suppressors of cdc25p overexpression identify two pathways that
RT   influence the G2/M checkpoint in fission yeast.";
RL   Genetics 150:1361-1375(1998).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- FUNCTION: Positively involved in the initiation of protein
CC       synthesis. Probable ATP-binding RNA helicase. Essential for cell
CC       viability. May play a role in mRNA splicing. Inactivation of ded1
CC       blocks mitotic cell cycle progression at G1 and G2/M.
CC   -!- SUBUNIT: Interacts with chk1, which is required for cell cycle
CC       arrest following DNA damage.
CC   -!- MISCELLANEOUS: A different form of ded1 has been identified via
CC       SDS-PAGE studies. It is uncertain how this form arises in vivo.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF025536; AAC04893.1; -.
DR   EMBL; AF084222; AAC34121.1; -.
DR   EMBL; AB012389; BAA25324.1; -.
DR   EMBL; AJ237697; CAB40192.1; -.
DR   EMBL; AL022598; CAA18646.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_ATP_helicase.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; Cell cycle; Mitosis.
FT   NP_BIND     213    220       ATP (POTENTIAL).
FT   SITE        336    339       DEAD BOX.
FT   CONFLICT     41     41       A -> R (IN REF. 3).
SQ   SEQUENCE   636 AA;  69758 MW;  094630A41A3C26F1 CRC64;
```

```
Alignment Scores:
Pred. No.:               1.12e-70       Length:          636
Score:                   1109.00        Matches:         271
Percent Similarity:      56.87%         Conservative:    89
Best Local Similarity:   42.81%         Mismatches:      185
Query Match:             28.05%         Indels:          88
DB:                      1              Gaps:            17
```

US-09-714-865-15 (1-2172) x DED1_SCHPO (1-636)

```
Qy        382 TCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGA 441
              |||   ||||||       |||               ||||||   |||      |||
Db         54 SerSerArgGlySerSerArgSerGlnThrProSerGluPheSerSerAsnTyr------ 71

Qy        442 GGTGGAAGAGGTAGTTTC---CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAAT 498
              |||||||||      :::   ||||||          |||   |||   |||      |||
Db         72 GlyGlyArgArgGluTyrAsnArgGlyGlyHisTyrGlyGlyGlyGluGlyArgGlnAsn 91

Qy        499 AATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGA 558
              |||                                      ||||||       |||
Db         92 Asn---------------------------TyrArgGlyGlyArgGluGly--------- 99

Qy        559 CCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT 618
```

```
                           |||      ||||||         ::::::        |||   |||
Db       100 ------------GlyTyrSerAsnGlyGlyGlyTyrArgAsnAsnArgGlyPheGlyGln 115

Qy       619 GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT 678
                |||   |||                  :::   ||||||          :::|||:::
Db       116 TrpArgAspGly---------------GlnHisValIle------GlyAlaArgAsnThr 128

Qy       679 TGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTAC 738
                                 |||    :::|||        |||   |||||||:::
Db       129 LeuLeuGluArgGlnLeuPheGlyAlaValAlaAsp------GlyThrLysValSer--- 145

Qy       739 ATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATA 798
                                                                 |||||||||
Db       146 ---------------------------------------------------ThrGlyIle 148

Qy       799 AACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCA 858
             ||||||:::|||||||||    |||   ||||||||||||||    |||        |||
Db       149 AsnPheGluLysTyrAspAspIleProValGluValSerGlyGlyAspIleGluPro--- 167

Qy       859 ATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCT 918
             :::        |||            |||          |||      ||||||      :::
Db       168 ValAsnGluPheThrSerProProLeuAsnSerHisLeuLeuGlnAsnIleLysLeuSer 187

Qy       919 GGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGA 978
             |||||||||:::   |||||||||||||||    ||||||||||||:::    :::||||||
Db       188 GlyTyrThrGlnProThrProValGlnLysAsnSerIleProIleValThrSerGlyArg 207

Qy       979 GATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
             ||||||||||||||||||||||||||||||||||||||||||||||||    ||||||    ||||||
Db       208 AspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaGlyPheLeuPheProIle 227

Qy      1039 TTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG------------ 1086
             |||:::             |||      |||            :::
Db       228 LeuSerLeuAlaPheAspLysGlyProAlaAlaValProValAspGlnAspAlaGlyMet 247

Qy      1087 ---------------TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTG 1131
                                   |||         :::|||:::||||||||||||||||||
Db       248 GlyTyrArgProArgLysAlaTyrProThrThrLeuIleLeuAlaProThrArgGluLeu 267

Qy      1132 GTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTT 1191
             |||   ||||||:::   |||:::|||||||||       :::    :::    ||||||
Db       268 ValCysGlnIleHisGluGluSerArgLysPheCysTyrArgSerTrpValArgProCys 287

Qy      1192 GTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAAT 1251
                :::||||||||         :::            ||||||||||||   |||||||||:::
Db       288 AlaValTyrGlyGlyAlaAspIleArgAlaGlnIleArgGlnIleAspGlnGlyCysAsp 307

Qy      1252 ATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTC 1311
             :::|||   |||||||||||||||||:::|||:::|||      :::     :::|||    |||
Db       308 LeuLeuSerAlaThrProGlyArgLeuValAspLeuIleAspArgGlyArgIleSerLeu 327

Qy      1312 AAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGT 1371
                      ||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       328 AlaAsnIleLysPheLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlu 347

Qy      1372 CCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACC 1431
             |||:::::::::   ::::::                |||    |||    |||:::|||||||||
Db       348 ProGlnIleArgHisIleValGluGlyAlaAspMetThrSerValGluGluArgGlnThr 367

Qy      1432 CTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAG 1491
             |||||||||||||||||||||||||||     :::||||||    ||||||    :::||||||||||
Db       368 LeuMetPheSerAlaThrPheProArgAspIleGlnLeuLeuAlaArgAspPheLeuLys 387

Qy      1492 TCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAG 1551
                :::|||:::|||::::::||||||:::||||||              ::::::    |||
Db       388 ---AspTyrValPheLeuSerValGlyArgValGlySerThrSerGluAsnIleThrGln 406

Qy      1552 ACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611
                |||:::    |||                ||||||    |||::::::||||||
Db       407 LysValValHisValGluAspSerGluLysArgSerTyrLeuLeuAspIleLeuHisThr 426

Qy      1612 ATAGGGGATGAA---AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACT 1668
             :::         |||         |||::::::|||||||||||||:::         ||||||
Db       427 LeuProProGluGlyLeuThrLeuIlePheValGluThrLysArgMetAlaAspThrLeu 446

Qy      1669 GCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAG 1728
                :::|||                              ||||||||||||||||||   |||
Db       447 ThrAspTyrLeuLeuAsnSerAsnPheProAlaThrSerIleHisGlyAspArgThrGln 466

Qy      1729 AGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCT 1788
             ||||||||||||:::||||||          ||||||   |||:::          ::::::||||||
Db       467 ArgGluArgGluArgAlaLeuGluLeuPheArgSerGlyArgThrSerIleMetValAla 486

Qy      1789 ACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGAT 1848
             |||:::||||||:::|||||||||||||||||    ||||||    ||||||||||||:::|||
Db       487 ThrAlaValAlaSerArgGlyLeuAspIleProAsnValThrHisValIleAsnTyrAsp 506

Qy      1849 CTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAAT 1908
             ||||||:::    |||||||:::||||||||||||||||||||||||||||||    ||||||
Db       507 LeuProThrAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAlaGlyAsn 526

Qy      1909 ACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTA 1968
             ||||||:::|||::::::||||||:::          ::::::       :::|||:::    |||
Db       527 ThrGlyGlnAlaValAlaPhePheAsn---ArgAsnAsnLysGlyIleAlaLysGluLeu 545

Qy      1969 GTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTT 2028
             :::::::::|||    :::|||    |||:::    |||::::::|||          :::|||
Db       546 IleGluLeuLeuGlnGluAlaAsnGlnGluCysProSerPheLeuIleAlaMetAlaArg 565

Qy      2029 AGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGAT 2088
               :::                |||    |||:::    ||||||          ::::::
Db       566 GluSer---------SerPheGlyGlyAsnGlyArgGlyGlyArgTyrSerGlyArgGly 582

Qy      2089 ACCAGAAAGGGC-------------------------AAGAGCACTTTGAACACAGCT 2121
                |||   |||                         :::    |||    :::::::::
Db       583 GlyArgGlyGlyAsnAlaTyrGlyAlaArgAspPheArgArgProThrAsnSerSerSer 602

Qy      2122 GGGTTTTCTTCT---------------------------------TCACGAGCTCCC--- 2145
             |||:::||||||                                  ||||||    |||
Db       603 GlyTyrSerSerGlyProSerTyrSerGlyTyrGlyGlyPheGluSerArgThrProHis 622

Qy      2146 ---------------AATCCAGTAGATGATGAGTCATGG 2169
                             |||           :::||||||
Db       623 HisGlyAsnThrTyrAsnSerGlySerAlaGlnSerTrp 635
```

```
RESULT 12
DBP1_YEAST
ID   DBP1_YEAST     STANDARD;      PRT;   617 AA.
AC   P24784; P20446;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable ATP-dependent RNA helicase DBP1 (Helicase CA1).
GN   DBP1 OR YPL119C OR LPH8C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBY939;
RX   MEDLINE=91312117; PubMed=1857205;
RA   Jamieson D.J., Beggs J.D.;
RT   "A suppressor of yeast spp81/ded1 mutations encodes a very similar
RT   putative ATP-dependent RNA helicase.";
RL   Mol. Microbiol. 5:805-812(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313271; PubMed=9169875;
RA   Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
RA   Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA   Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA   Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA   Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA   Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
```

```
RA   Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA   Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA   Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA   Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA   Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA   Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA   Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA   Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA   Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA   Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL   Nature 387:103-105(1997).
RN   [3]
RP   SEQUENCE OF 316-500 FROM N.A.
RX   MEDLINE=90160368; PubMed=2406722;
RA   Chang T.-H., Arenas J., Abelson J.;
RT   "Identification of five putative yeast RNA helicase genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:1571-1575(1990).
CC   -!- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE. SUPPRESSOR OF A
CC       SPP81/DED1 MUTATION.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X55993; CAA39465.1; -.
DR   EMBL; U43503; AAB68243.1; -.
DR   PIR; S16790; S16790.
DR   PIR; A34848; A34848.
DR   HSSP; Q58083; 1HV8.
DR   SGD; S0006040; DBP1.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding.
FT   NP_BIND      198    205       ATP (BY SIMILARITY).
FT   SITE         318    321       DEAD BOX.
FT   CONFLICT      43     44       ST -> RS (IN REF. 1).
FT   CONFLICT      48     48       E -> K (IN REF. 1).
FT   CONFLICT      88     88       G -> R (IN REF. 1).
FT   CONFLICT     115    115       E -> QK (IN REF. 1).
FT   CONFLICT     496    496       V -> I (IN REF. 3).
SQ   SEQUENCE   617 AA;  67917 MW;  8152404B7628671D CRC64;

Alignment Scores:
Pred. No.:              1.31e-70      Length:        617
Score:                  1108.00       Matches:       252
Percent Similarity:     57.43%        Conservative:  92
Best Local Similarity:  42.07%        Mismatches:    185
Query Match:            28.02%        Indels:        70
DB:                     1             Gaps:          14

US-09-714-865-15 (1-2172) x DBP1_YEAST (1-617)

Qy        409 AATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC--------- 459
              ||||||  |||      |||             |||||:::  |||:::
Db         14 AsnAsnLysGluAsnGlyGly---------GlyGlyGlyLysSerSerTyrValProPro 30

Qy        460 ------------CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTA 507
                          ||||||  :::      |||          :::|||      :::
Db         31 HisLeuArgSerArgGly---LysProSerPheGluArgSerThrProLysGlnGlu--- 48

Qy        508 GACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTA 567
                      |||:::            |||||||||   |||      |||
Db         49 ------AspLysVal---------ThrGlyGlyAspPhePheArgArg------------ 59

Qy        568 AGTGGC------ACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAA 621
              :::|||      |||||||||
Db         60 AlaGlyArgGlnThrGlyAsn--------------------------------------- 66

Qy        622 CGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGG 681
                 ||||||:::   |||   ::::::|||           |||:::      |||
Db         67 AsnGlyGlyPhePheGlyPheSerLysGluArgAsnGlyGlyThrSerAlaAsnTyrAsn 86

Qy        682 AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATA 741
              :::      :::         :::||||||      :::               :::|||
Db         87 ArgGlyGlySerSerAsnTyrLysSerSerGlyAsnArgTrpValAsnGlyLysHisIle 106

Qy        742 CCC---CCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA------------------ 780
              |||   |||         :::   |||      :::|||
Db        107 ProGlyProLysAsnAlaLysLeuGluAlaGluLeuPheGlyValHisAspAspProAsp 126

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
              ::::::   :::||||||   ||||||   ||||||   |||   |||:::   ||||||
Db        127 TyrHisSerSerGlyIleLysPheAspAsnTyrAspAsnIleProValAspAlaSerGly 146

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
                 |||   |||      ||||||   |||               |||   :::   |||
Db        147 LysAspValProGluProIleLeuAspPheSerSerProProLeuAspGluLeuLeuMet 166

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
                 ||||||      |||   :::||||||   ||||||||||||||||||||||||
Db        167 GluAsnIleLysLeuAlaSerPheThrLysProThrProValGlnLysTyrSerIlePro 186

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              |||:::      |||||||||||||||||||||||||||||||||||||||||||
Db        187 IleValThrLysGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrGly 206

Qy       1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACT------------ 1068
                 ||||||   |||:::         :::         |||   :::
Db        207 GlyPheLeuPheProLeuPheThrGluLeuPheArgSerGlyProSerProValProGlu 226

Qy       1069 ------GCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACT 1122
                    |||   :::               |||      :::::::||||||||
Db        227 LysAlaGlnSerPheTyrSerArgLysGlyTyrProSerAlaLeuValLeuAlaProThr 246

Qy       1123 CGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTA 1182
              |||||||||      ||||||:::   ||||||||||||||::::::   :::   |||
Db        247 ArgGluLeuAlaThrGlnIlePheGluGluAlaArgLysPheThrTyrArgSerTrpVal 266

Qy       1183 AGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAA 1242
              |||      |||:::|||||||||         :::|||:::   :::|||::::::   :::
Db        267 ArgProCysValValTyrGlyGlyAlaProIleGlyAsnGlnMetArgGluValAspArg 286

Qy       1243 GGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAG 1302
              ||||||::::::|||   ||||||||||||||||||   |||::::::   :::   |||
Db        287 GlyCysAspLeuLeuValAlaThrProGlyArgLeuAsnAspLeuLeuGluArgGlyLys 306

Qy       1303 ATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATG 1362
              :::   |||   |||||||||||||||||||||||||||||||||||||||||||||
Db        307 ValSerLeuAlaAsnIleLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMet 326

Qy       1363 GGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAG 1422
              ||||||   |||:::::::::   ::::::               |||||||||   |||
Db        327 GlyPheGluProGlnIleArgHisIleValGluGluCysAspMetProSerValGluAsn 346

Qy       1423 CGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAG 1482
              ||||||||||||||||||||||||||||||   :::||||||   ||||||   :::
Db        347 ArgGlnThrLeuMetPheSerAlaThrPheProValAspIleGlnHisLeuAlaArgAsp 366

Qy       1483 TTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGAT 1542
              ||||||         ||||||:::|||::::::||||||:::||||||         :::
Db        367 PheLeu---AspAsnTyrIlePheLeuSerValGlyArgValGlySerThrSerGluAsn 385

Qy       1543 GTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATT 1602
```

```
                 :::   |||   :::|||   |||             |||:::      |||:::::::::
Db        386 IleThrGlnArgIleLeuTyrValAspAspMetAspLysLysSerAlaLeuLeuAspLeu 405

Qy       1603 CTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGAT 1662
              |||                     |||::::::||||||||||||||:::   ||||||
Db        406 LeuSerAlaGluHisLysGlyLeuThrLeuIlePheValGluThrLysArgMetAlaAsp 425

Qy       1663 TTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGG 1722
                          ||||||         :::         |||:::|||||||||||||||
Db        426 GlnLeuThrAspPheLeuIleMetGlnAsnPheLysAlaThrAlaIleHisGlyAspArg 445

Qy       1723 GAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTT 1782
                 |||   |||||||||:::||||||      |||:::               :::|||
Db        446 ThrGlnAlaGluArgGluArgAlaLeuSerAlaPheLysAlaAsnValAlaAspIleLeu 465

Qy       1783 GTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAAT 1842
              |||||||||:::|||||||||||||||||||||||||   ||||||   ||||||||||||
Db        466 ValAlaThrAlaValAlaAlaArgGlyLeuAspIleProAsnValThrHisValIleAsn 485

Qy       1843 TTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGT 1902
              :::||||||||||||   ||||||:::|||||||||||||||||||||||||||||
Db        486 TyrAspLeuProSerAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAla 505

Qy       1903 GGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAG 1962
              ||||||||||||   |||   |||||||||:::      :::   |||   :::   :::
Db        506 GlyAsnThrGlyValAlaThrSerPhePheAsnSerAsnAsnGlnAsn---IleValLys 524

Qy       1963 CCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATT 2022
                 |||:::::::::|||   :::|||   |||:::||||||   :::|||   ::::::
Db        525 GlyLeuMetGluIleLeuAsnGluAlaAsnGlnGluValProThrPheLeuSerAspLeu 544

Qy       2023 GCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCA 2082
                                                                       |||
Db        545 ---------------------------------------------------------Ser 545

Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
                       :::|||      |||            ||||||   :::||||||
Db        546 ArgGlnAsnSerArgGlyGlyArgThrArgGlyGlyGlyGlyPhePheAsnSerArg 564
```

```
RESULT 13
GLH1_CAEEL
ID   GLH1_CAEEL     STANDARD;      PRT;   763 AA.
AC   P34689; Q9TXH4;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ATP-dependent RNA helicase glh-1 (Germline helicase-1).
GN   GLH-1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94022363; PubMed=8415696;
RA   Roussell D.L., Bennett K.L.;
RT   "glh-1, a germ-line putative RNA helicase from Caenorhabditis, has
RT   four zinc fingers.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:9300-9304(1993).
RN   [2]
RP   REVISIONS TO 83-138; 275; 288 AND 398.
RA   Roussell D.L., McCrone J.S., Smith P.A., Gruidl M.E., Bennett K.L.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.
CC   -!- DEVELOPMENTAL STAGE: DURING GERM-LINE PROLIFERATION.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 4 CCHC-TYPE ZINC FINGERS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L19948; AAC27384.1; -.
DR   PIR; A48686; A48686.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00098; zf-CCHC; 4.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 4.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS50158; ZF_CCHC; 4.
KW   Helicase; ATP-binding; RNA-binding; Zinc-finger; Repeat.
FT   DOMAIN       24     93       7 X 10 AA TANDEM REPEATS, GLY-RICH.
FT   REPEAT       24     33       1.
FT   REPEAT       34     43       2.
FT   REPEAT       44     53       3.
FT   REPEAT       54     63       4.
FT   REPEAT       64     73       5.
FT   REPEAT       74     83       6.
FT   REPEAT       84     93       7.
FT   DOMAIN      207    236       GLY-RICH.
FT   ZN_FING     158    175       CCHC-TYPE 1.
FT   ZN_FING     183    200       CCHC-TYPE 2.
FT   ZN_FING     242    259       CCHC-TYPE 3.
FT   ZN_FING     262    279       CCHC-TYPE 4.
FT   NP_BIND     385    392       ATP (BY SIMILARITY).
FT   SITE        499    502       DEAD BOX.
SQ   SEQUENCE   763 AA;  79792 MW;  ADB69DE286A028D6 CRC64;
```

```
Alignment Scores:
Pred. No.:               8.99e-70      Length:         763
Score:                   1096.50       Matches:        269
Percent Similarity:      48.89%        Conservative:   107
Best Local Similarity:   34.98%        Mismatches:     254
Query Match:             27.73%        Indels:         139
DB:                      1             Gaps:           17

US-09-714-865-15 (1-2172) x GLH1_CAEEL (1-763)

Qy        169 AGTGGATTTGCCTCTGGGCGGAAT------------TTTGGAAACAGAGATGCTGGTGAG 216
              |||||||||      |||:::|||            ||||||      :::   |||
Db         32 SerGlyPheGlyGlyGlyLysAsnGlyGlyThrGlyPheGlyGlyGlyAsnThrGlyGly 51

Qy        217 TGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGT------ 270
                                    :::         |||||||||   ||||||:::
Db         52 SerGlyPheGlyGlyGlyAsnThrGlyGlySerGlyPheGlyGlyGlyLysThrGlyGly 71

Qy        271 ------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT--------- 315
                    ||||||                  |||   |||      |||
Db         72 SerGlyPheGlyGlyGlyAsnThrCysGlySerGlyPheGlyGlyGlySerThrGlyGly 91

Qy        316 ------------------AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGAT 357
                                ||||||||||||               ::::::
Db         92 SerProTyrGlyGlyAlaSerSerGlyPheGlyGlySerThrAlaThrSerGlyPheGly 111

Qy        358 AATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCA 417
              :::      ::::::      |||         ||||||:::         |||   ::::::
Db        112 SerGlyGluLysSerSerAlaPheGlyGlySerGlyGlyPhe------GlyGlySerAla 129

Qy        418 GAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGA 477
                       |||                  |||   |||||||||   |||         |||
Db        130 ThrGlyPheGly------------SerGlyGlyGlySerPheGlyGlyGlyAsnSerGly 145

Qy        478 TTTGGTCTAGGA------------------------------------------------ 489
              ||||||   |||
```

```
Db        146 PheGlyGluGlyGlyHisGlyGlyGlyGluArgAsnAsnAsnCysPheAsnCysGlnGln 165

Qy        490 ---------------------------AGTCCAAATAATGACTTAGACCCAGACGAATGT 522
                                             |||    :::    :::|||      |||
Db        166 ProGlyHisArgSerSerAspCysProGluProArgLysGluArgGluProArgValCys 185

Qy        523 ATGCAG------------------------------------------------------ 528

Db        186 TyrAsnCysGlnGlnProGlyHisThrSerArgGluCysThrGluGluArgLysProArg 205

Qy        529 ------CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAAT 582
                    |||||||||||      |||                :::|||        :::
Db        206 GluGlyArgThrGlyGlyPheGlyGlyGlyAlaGlyPheGlyAsnAsnGlyGlyAsnAsp 225

Qy        583 GGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTA 642
              |||                     |||   |||   ||||||||     |||
Db        226 GlyPheGlyGlyAspGlyGlyPheGlyGlyGlyGluGluArgGlyProMetLysCysPhe 245

Qy        643 AAT--------------------------------------------------------- 645
              |||
Db        246 AsnCysLysGlyGluGlyHisArgSerAlaGluCysProGluProProArgGlyCysPhe 265

Qy        646 ---------GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAA 696
                       |||:::            |||         |||         :::   |||:::
Db        266 AsnCysGlyGluGln---------GlyHisArgSerAsnGluCysProAsnProAlaLys 282

Qy        697 GGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCT 756
                       |||    :::   :::|||||||||   |||||:::|||
Db        283 ProArgGluGlyValGluGlyGluGlyProLysAlaThrTyrValProValGluAspAsn 302

Qy        757 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGAC 816
                 ::::::                                  |||:::   |||:::|||:::
Db        303 MetGluAspValPheAsnMetGlnLysIleSerGluGlyLeuMetPheAsnLysPhePhe 322

Qy        817 ACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATT---------CTGACT 867
                          |||:::::::::         :::          |||         |||
Db        323 AspAlaGluValLysLeuThrSerSerGluLysThrValGlyIleLysProCysLysThr 342

Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
              |||   |||||||||||   :::|||:::      |||:::|||   |||||||||:::
Db        343 PheAlaGluAlaAsnLeuThrGluThrMetGlnLysAsnValAlaHisAlaGlyTyrSer 362

Qy        928 AAGCTTACTCCTGTGCAAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
              |||.   ||||||:::|||:::||||||:::::::::|||         |||      |||:::|||
Db        363 LysThrThrProIleGlnGlnTyrAlaLeuProLeuValHisGlnGlyTyrAspIleMet 382

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        383 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleMetThrArg 402

Qy       1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
              ::::::   |||            :::      :::          |||   ||||||
Db        403 LeuIleAspAspAsnAsnLeuAsnThrAlaGlyGluGlyGlyCysTyrProArgCysIle 422

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
              |||:::   |||||||||||||||   :::|||||||||   |||   |||||||||:::
Db        423 IleLeuThrProThrArgGluLeuAlaAspGlnIleTyrAsnGluGlyArgLysPheAla 442

Qy       1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
              :::   |||   :::               :::||||||||||      :::|||::::::
Db        443 TyrGlnThrMetMetGluIleLysProValTyrGlyGlyLeuAlaValGlyTyrAsnLys 462

Qy       1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
                 ||||||   :::|||      |||:::      |||   ||||||:::
Db        463 GlyGlnIleGluLysGlyAlaThrIleIleValGlyThrValGlyArgIleLysHisPhe 482

Qy       1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
                 :::      |||   |||   :::   ::::::   ||||||||||||||||||
Db        483 CysGluGluGlyThrIleLysLeuAspLysCysArgPhePheValLeuAspGluAlaAsp 502

Qy       1348 CGCATGTTGGAT---ATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
              ||||||:::|||   |||||||||||||   :::::::::   :::::::::
Db        503 ArgMetIleAspAlaMetGlyPheGlyThrAspIleGluThrIleValAsnTyrAspSer 522

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
              ||||||   ||||||   |||||||||||||||||||||||||||||||:::      :::
Db        523 MetProArgLysGluAsnArgGlnThrLeuMetPheSerAlaThrPheProAspSerVal 542

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
              |||      |||      ||||||:::   ||||||:::   :::|||:::   ::::::
Db        543 GlnGluAlaAlaArgAlaPheLeuArgGluAsnTyrValMetIleAlaIleAspLysIle 562

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
              |||   |||   :::   |||   |||               :::      :::   |||:::
Db        563 GlyAlaAlaAsnLysCysValLeuGlnGluPheGluArgCysGluArgSerGluLysLys 582

Qy       1585 GAAAAGCTCGTTGAAATTCTG-----------------------------------CGA 1608
              :::||||||:::|||:::|||
Db        583 AspLysLeuLeuGluLeuLeuGlyIleAspIleAspSerTyrThrThrGluLysSerAla 602

Qy       1609 AACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACT 1668
                 :::      :::::::::||||||||||||||         :::         ||||||
Db        603 GluValTyrThrLysLysThrMetValPheValSerGlnArgAlaMetAlaAspIleLeu 622

Qy       1669 GCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAG 1728
              |||:::   |||         ::::::         :::||||||||   ||||||||
Db        623 AlaSerIleLeuSerSerAlaGlnValProAlaIleThrIleHisGlyAlaArgGluGln 642

Qy       1729 AGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCT 1788
              |||||||||   :::||||||      ||||||   |||      |||||||||:::|||
Db        643 ArgGluArgSerGluAlaLeuArgGlnPheArgAsnGlySerLysProValLeuIleAla 662

Qy       1789 ACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGAT 1848
              |||:::||||||   |||||||||||||||:::   |||   ||||||||||||:::|||
Db        663 ThrAlaValAlaGluArgGlyLeuAspIleLysGlyValAspHisValIleAsnTyrAsp 682

Qy       1849 CTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAAT 1908
              :::|||      ||||||:::|||:::||||||||||||||||||||||||   ||||||
Db        683 MetProAspAsnIleAspAspTyrIleHisArgIleGlyArgThrGlyArgValGlyAsn 702

Qy       1909 ACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTA 1968
              :::|||||||||   ||||||         :::      |||      |||      |||
Db        703 SerGlyArgAlaThrSerPhe------IleSerGluAspCysSerLeuLeuSerGluLeu 720

Qy       1969 GTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTT 2028
              |||   ||||||   ||||||||||||   ||||||   |||::::::
Db        721 ValGlyValLeuAlaAspAlaGlnGlnIleValProAspTrpMetGln------------ 736

Qy       2029 AGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGAT 2088
                                      |||:::      ||||||      ||||||
Db        737 ------------------------GlyAlaAlaGlyGlyAsnTyrGlyAlaSer------ 746

Qy       2089 ACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAAT 2148
                                              ||||||   ||||||      |||
Db        747 --------------------------------GlyPheGlySerSer---ValProThr 754

Qy       2149 ------CCAGTAGATGATGAGTCATGG 2169
                    |||   |||:::|||   |||
Db        755 GlnValProGlnAspGluGluGlyTrp 763
```

```
RESULT 14
GLH3_CAEEL
ID   GLH3_CAEEL     STANDARD;      PRT;   720 AA.
AC   O01836;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ATP-dependent RNA helicase glh-3 (Germline helicase-3).
GN   GLH-3 OR B0414.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
```

```
RX   MEDLINE=20311358; PubMed=10851135;
RA   Kuznicki K.A., Smith P.A., Leung-Chiu W.M., Estevez A.O., Scott H.C.,
RA   Bennett K.L.;
RT   "Combinatorial RNA interference indicates GLH-4 can compensate for
RT   GLH-1; these two P granule components are critical for fertility in
RT   C. elegans.";
RL   Development 127:2907-2916(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Sammons L., Wohldmann P., Rohlfing T.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.
CC   -!- DEVELOPMENTAL STAGE: DURING GERM-LINE PROLIFERATION.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF079509; AAC28388.1; -.
DR   EMBL; AF003145; AAB57719.1; -.
DR   HSSP; Q58083; 1HV8.
DR   WormPep; B0414.6; CE07736.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00098; zf-CCHC; 2.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 2.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS50158; ZF_CCHC; 2.
KW   Helicase; ATP-binding; RNA-binding; Zinc-finger; Repeat.
FT   ZN_FING      202    219       CCHC-TYPE 1.
FT   ZN_FING      222    239       CCHC-TYPE 2.
FT   NP_BIND      342    349       ATP (BY SIMILARITY).
FT   SITE         456    459       DEAD BOX.
SQ   SEQUENCE   720 AA;  79728 MW;  A91F49FE2B79DEAA CRC64;

Alignment Scores:
Pred. No.:               1.94e-62     Length:          720
Score:                   992.50       Matches:         248
Percent Similarity:      51.20%       Conservative:    136
Best Local Similarity:   33.07%       Mismatches:      237
Query Match:             25.10%       Indels:          129
DB:                      1            Gaps:            22

US-09-714-865-15 (1-2172) x GLH3_CAEEL (1-720)

Qy        130 ATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGG 189
              |||||     |||:::::            :::::::::  ||||||        :::
Db          1 MetAspLysSerProThrLysThrSer------IleArgThrLysPheAlaArgHisGln 18

Qy        190 AATTTTGGAAACAGAGATGCTGGTGAG------------TGTAATAAGCGAGATAATACA 237
                   :::   |||      |||            |||   |||:::|||:::
Db         19 ProIleSerAspValAspThrThrGluGlnSerSerSerCysIleLysLysAspAspArg 38

Qy        238 TCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGA--------------- 282
                   :::     ||||||||      |||      :::|||
Db         39 Gly---LeuSerSerPheGlyValGlnSerSerValPheSerArgArgSerCysArgMet 57

Qy        282 ------------------------------------------------------------ 282
               
Db         58 SerGluLeuGluAlaLysProThrIleIleSerGluAspGlnArgIleAlaValArgSer 77

Qy        283 ------GGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAG 336
                    |||   |||   |||   |||:::|||         :::   |||
Db         78 GluIleGlyGlySerPheSerGlyPheAspAspLysValAspAsnValPhe------His 95

Qy        337 TCTAGTAATGACTGCGAAGATAATCCA--------------------------------- 363
              |||:::|||:::          :::|||
Db         96 SerAsnAsnAsnLeuHisGlySerProSerThrThrGluLeuGluCysProGlyIleMet 115

Qy        364 ---------------------------------ACACGGAACAGAGGGTTTTCCAAGAGA 390
                                               :::|||:::|||     :::   |||
Db        116 AsnProArgPheLeuValGlyArgSerLeuAsnSerArgSerArgAlaValThr---Arg 134

Qy        391 GGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCA---TACAGAAGAGGTGGA 447
               |||      |||      |||      |||      |||      :::|||
Db        135 GlySerLysArgThrSerAsnValLysGluAsnGluGlySerIleHisArgSerAspAsp 154

Qy        448 AGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTA 507
               :::    |||         |||                  :::   :::   :::
Db        155 GlnValSerThrGluAsnCys------------------SerAlaLysAspGluGluArg 168

Qy        508 GACCCAGACGAATGTATGCAGCGCACTGGTGGCCTT------TTTGGTTCTAGAAGA--- 558
              |||   |||                  :::||||||:::      :::|||::::::|||
Db        169 AspArgAsp---------------SerGlyGlyValSerSerTyrGlyAsnLysArgSer 183

Qy        559 ---------------------CCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAA 597
                                   |||:::|||         |||   |||
Db        184 AspGluPheCysGlyThrSerProIleLeuGluAlaLysGlyPheGly---------Ile 200

Qy        598 AGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTA--------------- 642
              |||   :::             ::::::   |||::::::
Db        201 SerAsnThrCysPheAsnCysLysLysTyrGlyHisArgAlaThrGluCysSerAlaPro 220

Qy        643 AATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGA 702
                    |||                     |||         :::|||
Db        221 GlnArgGluCysAlaAsnCysGlyAspProAsnHisArgAlaAsnGluCysAlaSerTrp 240

Qy        703 GAAAGTAGTGATACTCAAGGACCA---AAAGTGACCTACATACCCCCTCCTCCACCTGAG 759
                    :::      |||   |||   ||||||||||||:::|||
Db        241 SerLysAsnGlyValGlnGluProThrLysValThrTyrValProValValAspLysMet 260

Qy        760 GATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACT 819
              :::|||                           |||      |||||||||:::
Db        261 GluGluValPheSerMetLeuLysIleAsnAlaGlyAspPhePheAspLysPhePheAsp 280

Qy        820 ATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATT---------CTGACTTTT 870
                       |||::::::         |||      |||         :::|||
Db        281 AlaSerValGlnLeuValSerArgGlyGlnProValThrIleGlnProCysLysSerPhe 300

Qy        871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
                 ::::::::::::   |||::::::      |||:::   :::||||||||||||:::
Db        301 SerAspSerAspIleProGlnSerMetArgArgAsnValGluArgAlaGlyTyrThrArg 320

Qy        931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
                 ||||||:::|||:::|||::::::|||::::::      |||:::|||::::::|||
Db        321 ThrThrProIleGlnGlnTyrThrLeuProLeuValAlaAspGlyLysAspIleLeuAla 340

Qy        991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG 1050
              |||||||||||||||||||||||||||||||||||||||||||||||||::::::   :::
Db        341 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleMetSerArgLeu 360

Qy       1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAG------------------ 1092
              :::   :::                     |||:::|||
Db        361 IleLeuGlu---------------------LysAspLeuAsnTyrGlyAlaGluGlyGly 373

Qy       1093 ---GAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149
                 |||   |||||||||:::   |||||||||||||||   :::|||||||||
Db        374 CysTyrProArgCysIleIleLeuThrProThrArgGluLeuAlaAspGlnIleTyrAsn 393

Qy       1150 GAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGAACC 1209
              |||   ||||||||||||:::   :::   :::                 :::|||||||||
Db        394 GluGlyArgLysPheSerTyrGlnSerValMetGluIleLysProValTyrGlyGlyIle 413
```

```
Qy       1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269
                :::|||::::::          |||||:::::::|||||   |||:::      |||
Db        414 AsnValGlyTyrAsnLysSerGlnIleMetLysGlyCysThrIleIleValGlyThrIle 433

Qy       1270 GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA 1329
              ||||||:::                         |||   |||    :::    :::||||||
Db        434 GlyArgValLysHisPheCysGluAspGlyAlaIleLysLeuAspLysCysArgTyrLeu 453

Qy       1330 GTTTTGGATGAAGCTGATCGCATGTTGGAT---ATGGGTTTTGGTCCAGAAATGAAGAAG 1386
              ||||||||||||||||||||||||:::|||   |||||||||||||||||:::::::::
Db        454 ValLeuAspGluAlaAspArgMetIleAspSerMetGlyPheGlyProGluIleGluGln 473

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              :::|||:::             ||||||       ::::::|||||||||:::||||||||||
Db        474 IleIleAsnTyrLysAsnMetProLysAsnAspLysArgGlnThrMetMetPheSerAla 493

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              |||||||||       :::|||         |||    :::     |||:::     :::|||
Db        494 ThrPheProSerSerValGlnGluAlaAlaArgLysLeuLeuArgGluAspTyrThrMet 513

Qy       1507 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1566
              :::    :::    ::::::|||    |||    :::     |||     |||
Db        514 IleThrIleAspLysIleGlyAlaAlaAsnLysCysValIleGlnGluPheGluLeuCys 533

Qy       1567 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTG-------------------- 1605
                 :::    ||||||    :::||||||:::::::::|||
Db        534 AspArgThrSerLysValAspLysLeuLeuLysLeuLeuGlyIleAspIleAspThrTyr 553

Qy       1606 ---------------CGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAG 1650
                                   ::::::       ::::::|||:::|||||||||        :::
Db        554 ThrThrGluLysAsnSerAspValPheValLysLysThrIleValPheValAlaGlnGln 573

Qy       1651 AAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATC 1710
              |||    ||||||       |||:::     :::              ::::::        :::|||
Db        574 LysMetAlaAspThrLeuAlaSerIleMetSerAlaAlaGlnValProAlaIleThrIle 593

Qy       1711 CATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTCGCTTTGGAAAG 1770
              ||||||    |||||||||:::||||||              ||||||        ||||||    |||
Db        594 HisGlyAlaArgGluGlnLysGluArgSerAlaAlaLeuLysLeuPheArgSerGlyAla 613

Qy       1771 TGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAA 1830
                 |||||||||:::||||||:::|||         |||||||||||||||:::    |||
Db        614 LysProValLeuIleAlaThrAlaValValGluArgGlyLeuAspIleLysGlyValAsp 633

Qy       1831 CATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGT 1890
              ||||||||||||:::|||:::|||:::      ||||||:::|||:::|||||||||||||||
Db        634 HisValIleAsnTyrAspMetProAsnAsnIleAspAspTyrIleHisArgIleGlyArg 653

Qy       1891 ACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAAC 1950
              |||||||||    ||||||:::||||||||||         ||||||         |||      |||
Db        654 ThrGlyArgValGlyAsnSerGlyArgAlaThrSerPheIleSerLeuAlaAspAspVal 673

Qy       1951 CATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGG 2010
                 :::          ||||||:::    |||    ||||||:::|||    ||||||:::|||
Db        674 GlnIleLeuProGlnLeuValArgThrLeuAlaAspAlaGluGlnValValProSerTrp 693

Qy       2011 TTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAAC 2070
              ::::::|||    |||                         |||::::::       |||
Db        694 MetLysGluAlaAla-----------------------GlyGlyThrSerAsnProAsn 705

Qy       2071 GTGTTT---GCATCAGTTGATACCAGAAAG 2097
                 |||      |||:::||||||    :::
Db        706 LysPheGluLysSerIleAspThrGluGlu 715

RESULT 15
DBP2_SCHPO
ID   DBP2_SCHPO     STANDARD;      PRT;   550 AA.
AC   P24782;
DT   01-MAR-1992 (Rel. 21, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   P68-like protein.
GN   DBP2 OR SPBP8B7.16C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91141480; PubMed=1996094;
RA   Iggo R.D., Jamieson D.J., McNeill S.A., Southgate J., McPheat J.,
RA   Lane D.P.;
RT   "p68 RNA helicase: identification of a nucleolar form and cloning of
RT   related genes containing a conserved intron in yeasts.";
RL   Mol. Cell. Biol. 11:1326-1333(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Matviw H., Yu G., Young D.;
RL   Submitted (FEB-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- FUNCTION: NOT KNOWN. ESSENTIAL PROTEIN. MAY BE A PUTATIVE RNA
CC       HELICASE.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX5/DDX17
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X52648; CAA36873.1; -.
DR   EMBL; L11574; AAA35319.1; -.
DR   EMBL; AL032684; CAA21801.1; -.
DR   PIR; S14048; S14048.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
```

```
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; RNA-binding; Helicase; Nuclear protein.
FT   NP_BIND     166    173       ATP (BY SIMILARITY).
FT   SITE        276    279       DEAD BOX.
FT   DOMAIN      510    528       RNA-BINDING RGG-BOX.
FT   CONFLICT    235    235       R -> L (IN REF. 1).
FT   CONFLICT    240    240       R -> L (IN REF. 1).
SQ   SEQUENCE   550 AA;  61548 MW;  3BD1636D14275451 CRC64;

Alignment Scores:
Pred. No.:              2.51e-50      Length:        550
Score:                  820.50        Matches:       207
Percent Similarity:     51.96%        Conservative:  111
Best Local Similarity:  33.82%        Mismatches:    209
Query Match:            20.75%        Indels:        85
DB:                     1             Gaps:          21

US-09-714-865-15 (1-2172) x DBP2_SCHPO (1-550)

Qy        397 TATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGT 456
              |||||||||         :::|||  ||||||  |||        |||:::
Db          3 TyrArgAsp---------AsnGluTyrSerGlyAsnTyr------AsnGlyLysGluAsp 17

Qy        457 TTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGAC 516
                          ||||||  :::|||   |||     ||||||
Db         18 GlyTyrAsnSerArgGlyArgTyrGlyGlyGlyTyrArgAsnAsn--------------- 32

Qy        517 GAATGTATGCAGCGCACTGGTGGCCTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACA 576

Db         32 ------------------------------------------------------------ 32

Qy        577 GGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAA 636
                                ||||||   |||          |||||||||:::
Db         33 -----------------TyrSerArgGlyGlyGly---------ArgGlyGlyPheAsn 43

Qy        637 ---------------GGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGG 681
                             |||     ::::::        |||    |||:::|||     :::
Db         44 AspGlyAlaSerTyrGlyTyrAspGlnArg---------GlyGlnGlyArgAsnPheTyr 60

Qy        682 AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATA 741
              :::|||:::      |||                                |||    |||
Db         61 GluSerAspGlyProGlyAlaAsnLeuValLysLysAspTrpLysAsnGluThrLeuIle 80

Qy        742 CCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAAC 801
              |||              :::          :::       ::::::           |||
Db         81 Pro--------------PheGlnLysAspPheTyrLysGluHisGluAsnValArgAsn 95

Qy        802 TTCGACAAATACGACACTATTCTTGTGGAA-----------------------GTGTCT 837
                    :::   |||       :::     |||                       |||
Db         96 ------ArgSerAspAlaGluValThrGluTyrArgLysGluLysGluIleValValHis 113

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
              |||     :::      |||        :::      |||||||||||||||         :::
Db        114 GlyLeuAsnValProLysProValThrThrPheGluGluAlaGlyPheProAsnTyrVal 133

Qy        898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
                    :::   :::   |||:::          ||||||:::|||:::     :::
Db        134 LeuLysGluValLysGlnLeuGlyPheGluAlaProThrProIleGlnGlnGlnAlaTrp 153

Qy        958 CCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACT 1017
              |||:::    ::::::||||||||||:::::        :::      |||||||||||||||||
Db        154 ProMetAlaMetSerGlyArgAspMetValGlyIleSerAlaThrGlySerGlyLysThr 173

Qy       1018 GCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGT 1077
                 ::::::  ||||||   :::     |||:::              :::    :::
Db        174 LeuSerTyrCysLeuProAlaIleValHisIleAsnAlaGlnProLeuLeuSerProGly 193

Qy       1078 TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAAC 1137
                          |||     :::::::::||||||||||||||||||
Db        194 ------------AspGlyProIleValLeuValLeuAlaProThrArgGluLeuAlaVal 209

Qy       1138 CAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATA 1197
              ||||||          |||         ||||||             :::     :::|||       :::
Db        210 GlnIleGlnGlnGluCysThrLysPheGlyLysSerSerArgIleArgAsnThrCysVal 229

Qy       1198 TATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTA 1257
              |||||||||         |||      ||||||   :::::::::|||        |||
Db        230 TyrGlyGlyValProArgGlyProGlnIleArgAspLeuIleArgGlyValGluIleCys 249

Qy       1258 TGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAG 1317
                 ||||||||||||||||||:::|||::::::          |||       |||::::::
Db        250 IleAlaThrProGlyArgLeuLeuAspMetLeuAspSerAsnLysThrAsnLeuArgArg 269

Qy       1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAA 1377
              :::   |||||||||||||||||||||||||||||||||||||||||    |||:::
Db        270 ValThrTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGln 289

Qy       1378 ATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATG 1437
              ::::::|||::::::                     :::      |||||||||:::|||
Db        290 IleArgLysIleValAsp-----------GlnIleArgProAspArgGlnThrValMet 305

Qy       1438 TTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAAT 1497
              ||||||||||||:::|||:::|||:::||||||||||||   ::::::|||   ::::::
Db        306 PheSerAlaThrTrpProLysGluValGlnArgLeuAlaArgAspTyrLeu---AsnAsp 324

Qy       1498 TATCTGTTTGTTGCTGTTGGACAAGTGGGT---GGAGCATGTAGAGATGTTCAGCAGACC 1554
              |||:::   |||   ||||||   :::          |||     :::::::::|||
Db        325 TyrIleGlnValThrValGlySerLeuAspLeuAlaAlaSerHisAsnIleLysGlnIle 344

Qy       1555 GTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTC-----------GTTGAAATT 1602
              |||     |||         ||||||   :::|||             |||:::
Db        345 ValGluValValAspAsnAlaAspLysArgAlaArgLeuGlyLysAspIleGluGluVal 364

Qy       1603 CTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGAT 1662
              |||::::::      |||    :::   ::::::|||       ||||||:::   ||||||
Db        365 LeuLysAsp---ArgAspAsnLysValLeuIlePheThrGlyThrLysArgValAlaAsp 383

Qy       1663 TTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGG 1722
                    ||||||   |||:::                :::||||||||||||:::
Db        384 AspIleThrArgPheLeuArgGlnAspGlyTrpProAlaLeuAlaIleHisGlyAspLys 403

Qy       1723 GAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTT 1782
                 |||   ||||||:::        |||      :::||||||   ||||||   |||::::::
Db        404 AlaGlnAspGluArgAspTrpValLeuAsnGluPheArgThrGlyLysSerProIleMet 423

Qy       1783 GTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAAT 1842
              |||||||||   ||||||:::|||||||:::|||::::::     :::    ||||||   |||
Db        424 ValAlaThrAspValAlaSerArgGlyIleAspValLysGlyIleThrHisValPheAsn 443

Qy       1843 TTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGT 1902
              :::|||   |||           ::::::|||||||||||||||||||||||||||
Db        444 TyrAspPheProGlyAsnThrGluAspTyrValHisArgIleGlyArgThrGlyArgAla 463

Qy       1903 GGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAAC---CATTTAGCA 1959
              |||      |||   |||   ::::::|||         |||||||||        |||
Db        464 GlyAlaLysGlyThrAlaTyrThrTyrPhe------ThrSerAspAsnAlaLysGlnAla 481

Qy       1960 CAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAA 2019
              :::     ||||||   :::|||::::::|||:::|||||||:::         |||||||||
Db        482 ArgGluLeuValSerIleLeuSerGluAlaLysGlnAspIleAspProLysLeuGluGlu 501

Qy       2020 ATTGCC---TTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTT 2076
              :::|||   :::|||:::          |||      |||:::   |||            :::
Db        502 MetAlaArgTyrSerSer------GlyGlyArgGlyGlyAsnTyrArgArgGlyGlyTyr 519

Qy       2077 GCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAAC---ACAGCTGGGTTTTCTTCT 2133
                       |||:::|||             |||         ||||||:::
Db        520 GlyArgGlyGlyPheArgArgGlyGlyGlyTyrGlyAsnArgAsnArgGlyPheThrGly 539

Qy       2134 TCACGAGCTCCCAATCCAGTAGATGATGAGTCATGG 2169
              |||   :::      |||:::               |||
Db        540 SerAsnSerAla---ProLeuAlaArgSerArgTrp 550

Search completed: June 10, 2003, 16:39:39
```

Job time : 92 secs

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:         June 10, 2003, 14:51:45 ; Search time 61.5 Seconds
                                           (without alignments)
                                           9412.044 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  3954
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:      1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111757_25998/app_query.fasta_1.2311
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_76_@runat_05062003_111757_25998 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      A_Geneseq_101002:*
                1:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3823	96.7	724	22	AAE02417	Human vasa protein
2	3247.5	82.1	713	22	AAE02419	Rat vasa protein.
3	3174	80.3	722	22	AAE02418	Mouse vasa protein
4	1997	50.5	700	22	AAE02421	Danio reio vasa pr
5	1959.5	49.6	700	22	AAE02420	Frog vasa protein.
6	1290	32.6	661	22	ABB59954	Drosophila melanog
7	1274	32.2	661	22	AAE02422	Fruit fly vasa pro
8	1242.5	31.4	798	22	ABB64631	Drosophila melanog
9	1229	31.1	662	23	ABB57045	Mouse ischaemic co
10	1215	30.7	662	19	AAW81502	Dead Box X (DBX) g
11	1215	30.7	662	19	AAW81501	Dead Box X (DBX) g
12	1215	30.7	662	21	AAY96483	Human DBX1, an RNA
13	1182.5	29.9	660	19	AAW81503	Dead Box Y (DBY) g
14	1099	27.8	612	21	AAG17433	Arabidopsis thalia
15	1099	27.8	612	21	AAG20655	Arabidopsis thalia
16	1095.5	27.7	646	21	AAG17458	Arabidopsis thalia
17	892.5	22.6	421	21	AAG17434	Arabidopsis thalia
18	892.5	22.6	421	21	AAG20656	Arabidopsis thalia
19	888	22.5	460	21	AAG17459	Arabidopsis thalia
20	814	20.6	399	21	AAG17435	Arabidopsis thalia
21	814	20.6	399	21	AAG20657	Arabidopsis thalia
22	813.5	20.6	438	21	AAG17460	Arabidopsis thalia
23	796.5	20.1	811	22	ABB64859	Drosophila melanog
24	795.5	20.1	574	22	ABB67213	Drosophila melanog
25	795.5	20.1	577	22	ABB64981	Drosophila melanog
26	795.5	20.1	578	22	ABB65733	Drosophila melanog
27	795.5	20.1	578	22	ABB67212	Drosophila melanog
28	778.5	19.7	615	22	ABB44573	Mouse wound healin
29	777.5	19.7	614	22	ABB44571	Mouse wound healin
30	777.5	19.7	614	22	ABB44572	Human wound healin
31	774.5	19.6	945	22	ABB65231	Drosophila melanog
32	769	19.4	703	22	ABB63551	Drosophila melanog
33	761	19.2	500	22	AAE10164	RNA helicase conse
34	758	19.2	344	21	AAY96482	Human topoisomeras
35	709.5	17.9	648	20	AAY49636	Human sdp3.8 (HAGE
36	708	17.9	1275	20	AAY29084	Human CIF130 prote
37	704	17.8	1275	21	AAY44262	Human cell cycle r
38	701	17.7	135	22	AAU23109	Novel human enzyme
39	696	17.6	1224	22	ABB62191	Drosophila melanog
40	691.5	17.5	791	22	ABB62354	Drosophila melanog
41	681.5	17.2	709	22	AAB76859	Human lung tumour
42	681.5	17.2	709	23	AAU85514	Clone #18929 of lu
43	681	17.2	802	22	ABB71892	Drosophila melanog
44	675	17.1	628	21	AAB42879	Human ORFX ORF2643
45	675	17.1	635	20	AAY73847	Human prostate tum

ALIGNMENTS

```
RESULT 1
AAE02417
ID   AAE02417 standard; Protein; 724 AA.
XX
AC   AAE02417;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human vasa protein.
XX
KW   Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;
KW   ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;
KW   colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;
KW   skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;
KW   medulloblastoma; choriocarcinoma; squamous cell carcinoma; leukaemia;
KW   acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;
KW   osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
```

```
KW   teratoma; mediastinal; intracranial.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          1..318
FT                   /note= "Highly antigenic"
FT   Region          117..146
FT                   /label= Epitope #1
FT   Region          695..723
FT                   /label= Epitope #2
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
DR   N-PSDB; AAD06354.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 18; Page 53-54; 66pp; English.
XX
CC   The present sequence is human vasa protein that has germ cell specific
CC   expression and is believed to play a determinative role in gonad
CC   development. Germ cells are specialised to produce haploid gametes in
CC   multicellular organisms. Vasa is useful in the diagnosis or treatment
CC   of conditions characterised by its aberrant expression and/or the
CC   presence of its mutant forms. The conditions include cancers such as
CC   biliary tract, brain, breast, colon, ovarian, pancreas, prostate,
CC   colorectal, oral, liver, lung, skin, basocellular, testis, renal,
CC   thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,
CC   melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   724 AA;

Alignment Scores:
Pred. No.:              0            Length:         724
Score:                  3823.00      Matches:        724
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            96.69%       Indels:         0
DB:                     22           Gaps:           0

US-09-714-865-15 (1-2172) x AAE02417 (1-724)

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAspGluAspTrpGluAlaGluIleAsnProHisMetSerSerTyrValProIle 20

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 PheGluLysAspArgTyrSerGlyGluAsnGlyAspAsnPheAsnArgThrProAlaSer 40

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerSerGluMetAspAspGlyProSerArgArgAspHisPheMetLysSerGlyPheAla 60

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerGlyArgAsnPheGlyAsnArgAspAlaGlyGluCysAsnLysArgAspAsnThrSer 80

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrMetGlyGlyPheGlyValGlyLysSerPheGlyAsnArgGlyPheSerAsnSerArg 100

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheGluAspGlyAspSerSerGlyPheTrpArgGluSerSerAsnAspCysGluAspAsn 120

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ProThrArgAsnArgGlyPheSerLysArgGlyGlyTyrArgAspGlyAsnAsnSerGlu 140

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AlaSerGlyProTyrArgArgGlyGlyArgGlySerPheArgGlyCysArgGlyGlyPhe 160

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GlyLeuGlySerProAsnAsnAspLeuAspProAspGluCysMetGlnArgThrGlyGly 180

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LeuPheGlySerArgArgProValLeuSerGlyThrGlyAsnGlyAspThrSerGlnSer 200

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ArgSerGlySerGlySerGluArgGlyGlyTyrLysGlyLeuAsnGluGluValIleThr 220

Qy        661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyGluSerSerAspThrGln 240

Qy        721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePheAla 260

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSerGly 280

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 HisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeuAsn 300

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIlePro 320

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 IleIleLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 340

Qy       1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaPheLeuLeuProIleLeuAlaHisMetMetHisAspGlyIleThrAlaSerArgPhe 360

Qy       1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuValAsnGln 380

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaValValIleTyr 400

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      401 GlyGlyThrGlnLeuGlyHisSerIleArgGlnIleValGlnGlyCysAsnIleLeuCys 420

Qy     1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 AlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLysGlnIle 440

Qy     1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      441 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlyProGluMet 460

Qy     1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      461 LysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeuMetPhe 480

Qy     1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 SerAlaThrPheProGluGluIleGlnArgLeuAlaAlaGluPheLeuLysSerAsnTyr 500

Qy     1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      501 LeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThrValLeu 520

Qy     1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      521 GlnValGlyGlnPheSerLysArgGluLysLeuValGluIleLeuArgAsnIleGlyAsp 540

Qy     1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GluArgThrMetValPheValGluThrLysLysLysAlaAspPheThrAlaThrPheLeu 560

Qy     1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      561 CysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGluArgGlu 580

Qy     1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      581 GlnAlaLeuGlyAspPheArgPheGlyLysCysProValLeuValAlaThrSerValAla 600

Qy     1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 AlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAspLeuProSerThr 620

Qy     1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      621 IleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAla 640

Qy     1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      641 IleSerPhePheAspLeuGluSerAspAsnHisLeuAlaGlnProLeuValLysValLeu 660

Qy     1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 ThrAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerThrTyrIle 680

Qy     2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      681 ProGlyPheSerGlySerThrArgGlyAsnValPheAlaSerValAspThrArgLysGly 700

Qy     2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      701 LysSerThrLeuAsnThrAlaGlyPheSerSerSerArgAlaProAsnProValAspAsp 720

Qy     2161 GAGTCATGGGAT 2172
            ||||||||||||
Db      721 GluSerTrpAsp 724
```

```
RESULT 2
AAE02419
ID   AAE02419 standard; Protein; 713 AA.
XX
AC   AAE02419;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Rat vasa protein.
XX
KW   Rat; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Rattus norvegicus.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 56-58; 66pp; English.
XX
CC   The present sequence is rat vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   713 AA;

Alignment Scores:
Pred. No.:              4.11e-301     Length:        713
Score:                  3247.50       Matches:       625
Percent Similarity:     90.55%        Conservative:  36
Best Local Similarity:  85.62%        Mismatches:    46
Query Match:            82.13%        Indels:        23
DB:                     22            Gaps:          6

US-09-714-865-15 (1-2172) x AAE02419 (1-713)

Qy        1 ATGGGGGATGAAGATTGGGAAGCAGAAATC---AACCCTCATATGTCTTCCTATGTTCCC 57
            ||||||||||||||||||||||||||||||      ||||||:::|||||||||||||||
Db        1 MetGlyAspGluAspTrpGluAlaGluIleLeuLysProHisValSerSerTyrValPro 20

Qy       58 ATATTTGAGAAGGATAGGTAT---TCTGGAGAAAATGGAGACAATTTTAACAGGACTCCA 114
            :::||||||||||||:::|||   ||||||   |||||||||   ||||||||||||
Db       21 ValPheGluLysAspLysTyrSerSerGlyAlaAsnGlyAspThrPheAsnArgThrSer 40

Qy      115 GCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGA 174
```

```
Db        41 AlaSerSerSerGluMetGluAspGlyProSerGlyArgAspHisPheMetArgSerGly 60

Qy       175 TTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT 234
Db        61 PheSerSerGlyArgAsnLeuGlyAsnArgAspIleGlyGluSerSerLysArgGluThr 80

Qy       235 ACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAAC 294
Db        81 ThrSerThrThrGlyGlyPheGlyArgGlyLysGlyPheGlyAsnArgGlyPheLeuAsn 100

Qy       295 AGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAA 354
Db       101 AsnLysPheGluGluGlyAspSerSerGlyPheTrpLysGluSerThrAsnAspCysGlu 120

Qy       355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
Db       121 AspThrGlnThrArgSerArgGlyPheSerLysArgGlyGlyTyrProAspGlyAsnAsp 140

Qy       415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGA 474
Db       141 SerGluAlaSerGlyProPheArgArgGlyGlyArgAspSer------------------ 154

Qy       475 GGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACT 534
Db       155 ---------------------------GluTyrAspGlnAspGlnGlySerGlnArgGly 165

Qy       535 GGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCT 594
Db       166 GlyGlyLeuPheGlySerArgLysProAlaAlaSerAspSerGlySerGlyAspThrPhe 185

Qy       595 CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654
Db       186 GlnSerArgSerGlyAsnAla------ArgGlyAlaTyrLysGlyLeuAsnGluGluVal 203

Qy       655 ATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGAT 714
Db       204 ValThrGlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyGluSerSerAsp 223

Qy       715 ACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATC 774
Db       224 IleGlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIle 243

Qy       775 TTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTG 834
Db       244 PheAlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluVal 263

Qy       835 TCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACA 894
Db       264 SerGlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThr 283

Qy       895 CTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGT 954
Db       284 LeuAsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSer 303

Qy       955 ATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAG 1014
Db       304 IleProIleValLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLys 323

Qy      1015 ACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGT 1074
Db       324 ThrAlaAlaPheLeuLeuProIleLeuAlaHisMetMetArgAspGlyIleThrAlaSer 343

Qy      1075 CGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTC 1134
Db       344 ArgPheLysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuIle 363

Qy      1135 AACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTT 1194
Db       364 AsnGlnIleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaValVal 383

Qy      1195 ATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATA 1254
Db       384 IleTyrGlyGlyThrGlnPheGlyHisSerIleArgGlnIleValGlnGlyCysAsnIle 403

Qy      1255 TTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAA 1314
Db       404 LeuCysAlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLys 423

Qy      1315 CAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCA 1374
Db       424 GlnValLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlyPro 443

Qy      1375 GAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTT 1434
Db       444 GluMetLysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeu 463

Qy      1435 ATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCA 1494
Db       464 LeuPheSerAlaThrPheProGluGluIleGlnArgLeuAlaGlyGluPheLeuLysSer 483

Qy      1495 AATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACC 1554
Db       484 AsnTyrLeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnSer 503

Qy      1555 GTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATA 1614
Db       504 IleLeuGlnValGlyProValPheLysLysArgLysLeuValGluIleLeuArgAsnIle 523

Qy      1615 GGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
Db       524 GlyAspGluArgProMetValPheValGluThrLysLysLysAlaAspPheIleAlaThr 543

Qy      1675 TTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
Db       544 PheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGlu 563

Qy      1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
Db       564 ArgGluGlnAlaLeuGlyAspPheArgCysGlyLysCysProValLeuValAlaThrSer 583

Qy      1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
Db       584 ValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAsnLeuPro 603

Qy      1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
Db       604 SerThrIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGly 623

Qy      1915 AGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAA 1974
Db       624 ArgAlaIleSerPhePheAspThrGluSerAspAsnHisLeuAlaGlnProLeuValLys 643

Qy      1975 GTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACA 2034
Db       644 ValLeuSerAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerSer 663

Qy      2035 TACATT---CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACC 2091
Db       664 TyrAlaProProSerPheSerAsnSerThrArgGlyAlaValPheAlaSerPheAspThr 683

Qy      2092 AGAAAG---------GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
Db       684 ArgLysAsnPheGlnGlyLysAsnThrLeuAsnThrAlaGlyIleSerSerAlaGlnAla 703

Qy      2143 CCCAATCCAGTAGATGATGAGTCATGGGAT 2172
Db       704 ProAsnProValAspAspGluSerTrpAsp 713
```

RESULT 3
AAE02418
ID AAE02418 standard; Protein; 722 AA.
XX
AC AAE02418;
XX
DT 10-AUG-2001 (first entry)

```
XX
DE   Mouse vasa protein.
XX
KW   Mouse; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Mus musculus.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 54-56; 66pp; English.
XX
CC   The present sequence is mouse vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   722 AA;

Alignment Scores:
Pred. No.:               4.39e-294   Length:         722
Score:                   3174.00     Matches:        611
Percent Similarity:      90.48%      Conservative:   45
Best Local Similarity:   84.28%      Mismatches:     59
Query Match:             80.27%      Indels:         11
DB:                      22          Gaps:           7

US-09-714-865-15 (1-2172) x AAE02418 (1-722)

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATC---AACCCTCATATGTCTTCCTATGTTCCC 57
              ||||||||||||||||||||||||||||||      ||||||:::||||||||||||||
Db          1 MetGlyAspGluAspTrpGluAlaGluIleLeuLysProHisValSerSerTyrValPro 20

Qy         58 ATATTTGAGAAGGATAGGTAT---TCTGGAGAAAATGGAGACAATTTTAACAGGACTCCA 114
              :::||||||||||:::|||      ||||||   |||||||||    ||||||||||||
Db         21 ValPheGluLysAspLysTyrSerSerGlyAlaAsnGlyAspThrPheAsnArgThrSer 40

Qy        115 GCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGA 174
              |||   |||||||||||:::|||||||||||    ||||||    ||||||:::||||||
Db         41 Ala---SerSerGluMetGluAspGlyProSerGlyArgAspAspPheMetArgSerGly 59

Qy        175 TTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT 234
              |||   ||||||||:::   |||:::||||||   ||||||    :::|||::::::|||
Db         60 PheProSerGlyArgSerLeuGlySerArgAspIleGlyGluSerSerLysLysGluAsn 79

Qy        235 ACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAAC 294
              ||||||||   ||||||||||||   ||||||   ||||||||||||||||||   |||
Db         80 ThrSerThrThrGlyGlyPheGlyArgGlyLysGlyPheGlyAsnArgGlyPheLeuAsn 99

Qy        295 AGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAA 354
              ::::::||||||:::||||||||||||||||||||:::||||||:::||||||||||||
Db        100 AsnLysPheGluGluGlyAspSerSerGlyPheTrpLysGluSerAsnAsnAspCysGlu 119

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
              ||||||   ||||||:::||||||||||||||||||||||||   :::|||||||||:::
Db        120 AspAsnGlnThrArgSerArgGlyPheSerLysArgGlyGlyCysGlnAspGlyAsnAsp 139

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGA 474
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        140 SerGluAlaSerGlyProPheArgArgGlyGlyArgGlySerPheArgGlyCysArgGly 159

Qy        475 GGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACT 534
              |||||||||||||||   ||||||::::::   |||   |||:::         |||
Db        160 GlyPheGlyLeuGlyArgProAsnSerGluSerAspGlnAspGlnGlyThrGlnCysGly 179

Qy        535 GGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCT 594
              ||||||            :::|||      |||   :::||||||||||||||||
Db        180 GlyGlyPheLeuValLeuGlyLysProAlaAlaSerAspSerGlyAsnGlyAspThrTyr 199

Qy        595 CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654
              |||||||||||||||||||||         ||||||||||||||||||||||||||||||
Db        200 GlnSerArgSerGlySerGly------ArgGlyGlyTyrLysGlyLeuAsnGluGluVal 217

Qy        655 ATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGAT 714
              :::|||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db        218 ValThrGlySerGlyLysAsnSerTrpLysSerGluThrGluGlyGlyGluSerSerAsp 237

Qy        715 ACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATC 774
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 SerGlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIle 257

Qy        775 TTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 PheAlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluVal 277

Qy        835 TCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACA 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 SerGlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThr 297

Qy        895 CTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGT 954
              |||||||||||||||   |||||||||||||||||||||||||||||||||||||||:::
Db        298 LeuAsnAsnAsnIleArgLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrThr 317

Qy        955 ATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAG 1014
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 IleProIleValLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLys 337

Qy       1015 ACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGT 1074
              |||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        338 ThrAlaAlaPheLeuLeuProIleLeuAlaHisMetMetArgAspGlyIleThrAlaSer 357

Qy       1075 CGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTC 1134
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        358 ArgPheLysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuIle 377

Qy       1135 AACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTT 1194
              ||||||||||||||||||||||||||||||||||||||||||||||   :::||||||
Db        378 AsnGlnIleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValIleSerValVal 397

Qy       1195 ATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATA 1254
              ||||||||||||||||||   ||||||||||:::||||||||||||||||||||||||||
```

```
Db        398 IleTyrGlyGlyThrGlnPheGlyHisSerValArgGlnIleValGlnGlyCysAsnIle 417

Qy       1255 TTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAA 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 LeuCysAlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLys 437

Qy       1315 CAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCA 1374
              |||:::|||||||||||||||||||||||||||   ||||||||||||||||   |||
Db        438 GlnValLysTyrLeuValLeuAspGluAlaAspSerMetLeuAspMetGlyPheAlaPro 457

Qy       1375 GAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTT 1434
              |||:::|||||||||||||||||||||||||||||||||||||||||||   ||||||||
Db        458 GluIleLysLysLeuIleSerCysProGlyMetProSerLysGluGlnHisGlnThrLeu 477

Qy       1435 ATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCA 1494
              :::||||||||||||||||||||||||||||||||||||||||||   :::|||||||||
Db        478 LeuPheSerAlaThrPheProGluGluIleGlnArgLeuAlaGlyAspPheLeuLysSer 497

Qy       1495 AATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACC 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 AsnTyrLeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThr 517

Qy       1555 GTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATA 1614
              :::||||||||||||||:::   |||   :::   |||:::               ||||||
Db        518 IleLeuGlnValGlyGlnTyrGlnLysGluLysSerLeuLeuArgPheTyrGluAsnIle 537

Qy       1615 GGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        538 GlyAspGluArgThrMetValPheValGluThrLysLysLysAlaAspPheIleAlaThr 557

Qy       1675 TTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db        558 PheLeuCysGlnGluLysIleSerSerThrSerIleHisGlyAspArgGluGlnArgGlu 577

Qy       1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
              ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        578 ArgGluGlnAlaLeuGlyAspPheArgCysGlyLysCysProValLeuValAlaThrSer 597

Qy       1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAspLeuPro 617

Qy       1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 SerThrIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGly 637

Qy       1915 AGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAA 1974
              |||||||||||||||||||   :::|||||||||||||||||||||||||||||||||||
Db        638 ArgAlaIleSerPhePheAspThrAspSerAspAsnHisLeuAlaGlnProLeuValLys 657

Qy       1975 GTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACA 2034
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 ValLeuSerAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerThr 677

Qy       2035 TACATT---CCTGGCTTCAGTGGTAGTACAAGAGGAAAC---GTGTTTGCATCAGTTGAT 2088
              |||:::   |||   ||||||   |||||||||||         ||||||||||||||||
Db        678 TyrValProProSerPheSerSerSerThrArgGlyGlyAlaValPheAlaSerValAsp 697

Qy       2089 ACCAGAAA----------GGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACG 2138
              ||||||           |||:::   |||   |||:::|||        |||||||||||
Db        698 ThrArgLysAsnTyrGlnGlyLysAlaHisValGluTyrSerGlyAspPhePhePheThr 717

Qy       2139 AGCTCCCAATCCAGT 2153
              |||||||||||||||
Db        718 SerSerGlnSerSer 722
```

RESULT 4
AAE02421

```
ID   AAE02421 standard; Protein; 700 AA.
XX
AC   AAE02421;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Danio reio vasa protein.
XX
KW   Vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Danio reio.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 59-61; 66pp; English.
XX
CC   The present sequence is Danio reio vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   700 AA;
```

```
Alignment Scores:
Pred. No.:               1.38e-181      Length:        700
Score:                   1997.00        Matches:       423
Percent Similarity:      67.37%         Conservative:  89
Best Local Similarity:   55.66%         Mismatches:    148
Query Match:             50.51%         Indels:        100
DB:                      22             Gaps:          20
```

US-09-714-865-15 (1-2172) x AAE02421 (1-700)

```
Qy         10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAG 69
                :::|||||||||   :::   :::|||   :::|||
Db          2 AspAspTrpGluGluAspGlnSerProValValSerCys-----------------Ser 15

Qy         70 GATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAA 129
                    :::   |||         |||                     |||:::||||||
Db         16 SerGlyPheGlyGlyAlaGlyAsnAspLys-----------------SerAsnSerGlu 29

Qy        130 ATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGG 189
```

```
                 :::|||   |||   :::                           :::|||
Db          30 GlyThrGluGlySerSerTrpLys-----------------------MetThrGlyAsp 41

Qy         190 AATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGT 249
               :::|||   |||                                             |||
Db          42 SerPheArgGlyArg-----------------------------------------Gly 47

Qy         250 GGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCA------AACAGCAGGTTT 303
               |||   |||   |||      |||   ||||||||         |||
Db          48 GlyArgGly---GlySerArgGlyGlyArgGlyGlyPheSerGlyPheLysSerGluIle 66

Qy         304 GAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCA 363
               ::::::    |||   |||   |||            |||   |||
Db          67 AspGluAsnGlySerAspGlyGlyTrp------------AsnGlyGlyGluSerArgGly 82

Qy         364 ACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA--- 420
                           ||||||      |||||||||:::|||   |||:::      |||
Db          83 ArgGlyArgGlyGlyPhe------ArgGlyGlyPheArgSerGlySerArgAspGluAsn 100

Qy         421 ---------------------------GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGT 453
                                                |||      |||   |||||||||
Db         101 AspGluAsnGlyAsnAspAspGlyTrpLysGlyGlyGluSerArgGlyArgGlyArgGly 120

Qy         454 AGTTTC---CGAGGTTGCCGTGGAGGATTT------------------------------ 480
                |||      |||   ||||||||||||
Db         121 GlyPheGlyGlyGlyPheArgGlyGlyPheArgAspGlyGlyAsnGluAspThrGlyArg 140

Qy         481 ---GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG--------- 528
                  |||   |||      ||||||:::         ||||||      :::
Db         141 ArgGlyPheGlyArgGluAsnAsnGluAsnGlyAsnAspGluGlyGlyGluGlyArgGly 160

Qy         529 ---CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGT 585
                        ||||||      |||   |||                  |||   |||
Db         161 ArgGlyArgGlyGlyPheArgGlyGlyPheArg---------------AspGlyGlyGly 175

Qy         586 GATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAAT 645
               |||   |||      |||         |||      |||||||||::::::|||   |||
Db         176 AspGluSerGlyLysArg---------GlyPheGlyArgGlyGlyPheArgGlyArgAsn 192

Qy         646 GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
               |||||||||   :::               :::   |||   :::   ||||||   |||
Db         193 GluGluValPheSer---LysValThrThrAlaAspLysLeuAspGlnGluGlySerGlu 211

Qy         706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAG 765
               ::::::             ||||||||||||   |||:::||||||||||||||:::|||
Db         212 AsnAla---------GlyProLysValValTyrValProProProProProGluGluGlu 228

Qy         766 GACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTT 825
                  |||||||||:::||||||   ||||||||||||||||||||||||   ||||||
Db         229 SerSerIlePheSerHisTyrAlaThrGlyIleAsnPheAspLysTyrAspAspIleLeu 248

Qy         826 GTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTC 885
               |||:::|||||||||   :::   |||   ||||||:::||||||||||||||   |||
Db         249 ValAspValSerGlySerAsnProProLysAlaIleMetThrPheGluGluAlaGlyLeu 268

Qy         886 TGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAA 945
               |||   :::|||:::   |||::::::|||:::||||||   |||   ||||||||||||
Db         269 CysAspSerLeuSerLysAsnValSerLysSerGlyTyrValLysProThrProValGln 288

Qy         946 AAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGG 1005
               |||:::   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         289 LysHisGlyIleProIleIleSerAlaGlyArgAspLeuMetAlaCysAlaGlnThrGly 308

Qy        1006 TCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATA 1065
               |||||||||||||||||||||||||||||||||||         |||   ||||||:::
Db         309 SerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnArgPheMetThrAspGlyVal 328

Qy        1066 ACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGA 1125
                  ||||||:::|||   |||:::||||||||||||   ||||||||||||||||||||
Db         329 AlaAlaSerLysPheSerGluIleGlnGluProGluAlaIleIleValAlaProThrArg 348

Qy        1126 GAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGA 1185
               ||||||:::||||||||||||||||||||||||||||||::::::||||||||||||||
Db         349 GluLeuIleAsnGlnIleTyrLeuGluAlaArgLysPheAlaTyrGlyThrCysValArg 368

Qy        1186 GCTGTTGTTATATATGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGC 1245
                  ||||||:::|||||||||         |||::::::||||||::::::::::::|||
Db         369 ProValValValTyrGlyGlyIleAsnThrGlyTyrThrIleArgGluValLeuLysGly 388

Qy        1246 TGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATT 1305
               ||||||:::||||||||||||||||||||||||||||   |||:::||||||:::   ||||||
Db         389 CysAsnValLeuCysAlaThrProGlyArgLeuHisAspLeuIleGlyArgGlyLysIle 408

Qy        1306 GGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGT 1365
               ||||||   :::::::::|||||||||||||||||||||||||||||||||||||||||
Db         409 GlyLeuSerLysValArgTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGly 428

Qy        1366 TTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGC 1425
               |||   |||||||||:::||||||::::::   |||||||||||||||||||:::|||
Db         429 PheGluProGluMetArgLysLeuValAlaSerProGlyMetProSerLysGluLysArg 448

Qy        1426 CAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTT 1485
               ||||||||||||||||||||||:::||||||:::|||||||||:::||||||:::|||
Db         449 GlnThrLeuMetPheSerAlaThrTyrProGluAspIleGlnArgMetAlaAlaAspPhe 468

Qy        1486 TTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTT 1545
               ||||||   :::|||:::|||:::|||||||||   |||||||||||||||   ||||||
Db         469 LeuLysValAspTyrIlePheLeuAlaValGlyValValGlyGlyAlaCysSerAspVal 488

Qy        1546 CAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTG 1605
               :::||||||::::::||||||   |||:::|||||||||::::::|||:::|||:::|||
Db         489 GluGlnThrIleValGlnValAspGlnTyrSerLysArgAspGlnLeuLeuGluLeuLeu 508

Qy        1606 CGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTT 1665
               |||   |||:::||||||||||||||||||||||||||||||||||:::   |||||||||
Db         509 ArgAlaThrGlyAsnGluArgThrMetValPheValGluThrLysArgSerAlaAspPhe 528

Qy        1666 ACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAA 1725
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         529 IleAlaThrPheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGlu 548

Qy        1726 CAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTT 1785
               ||||||||||||||:::||||||   |||||||||   |||   ||||||||||||||
Db         549 GlnArgGluArgGluLysAlaLeuSerAspPheArgLeuGlyHisCysProValLeuVal 568

Qy        1786 GCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTT 1845
               |||||||||||||||||||||||||||||||||   ||||||||||||:::||||||
Db         569 AlaThrSerValAlaAlaArgGlyLeuAspIleGluGlnValGlnHisValValAsnPhe 588

Qy        1846 GATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGG 1905
               |||:::||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db         589 AspMetProSerSerIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGly 608

Qy        1906 AATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCT 1965
               ||||||||||||||:::|||||||||:::   |||||||||      ||||||:::
Db         609 AsnThrGlyArgAlaValSerPhePheAsnProGluSerAspThrProLeuAlaArgSer 628

Qy        1966 CTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCC 2025
               |||||||||||||||:::   |||||||||   ||||||   ||||||||||||:::|||
Db         629 LeuValLysValLeuSerGlyAlaGlnGlnValValProLysTrpLeuGluGluValAla 648

Qy        2026 TTTAGTACATACATTCCTGGCTTCAGTGGT---AGTACAAGAGGAAACGTGTTTGCATCA 2082
               ||||||   :::      |||   :::|||   :::   ||||||   ||||||||||||
Db         649 PheSerAlaHis------GlyThrThrGlyPheAsnProArgGlyLysValPheAlaSer 666

Qy        2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
               |||:::|||||||||   |||                  |||   |||
Db         667 ThrAspSerArgLysGlyGlySer------------------PheLysSerAspGluPro 680

Qy        2143 CCC-----------------------------AATCCAGTAGATGATGAGTCATGGGAT 2172
               |||                                         |||||||||   |||:::
Db         681 ProProSerGlnThrSerAlaProSerAlaAlaAlaAlaAlaAspAspGluGluTrpGlu 700
```

RESULT 5
AAE02420
ID AAE02420 standard; Protein; 700 AA.

```
XX
AC   AAE02420;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Frog vasa protein.
XX
KW   Frog; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial..
XX
OS   Xenopus laevis.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;    99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 58-59; 66pp; English.
XX
CC   The present sequence is frog vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   700 AA;

Alignment Scores:
Pred. No.:              5.29e-178      Length:         700
Score:                  1959.50        Matches:        409
Percent Similarity:     68.17%         Conservative:   105
Best Local Similarity:  54.24%         Mismatches:     153
Query Match:            49.56%         Indels:         87
DB:                     22             Gaps:           18

US-09-714-865-15 (1-2172) x AAE02420 (1-700)

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
                :::|||:::|||:::   |||||               :::|||||||||   |||
Db          2 GluGluAsnTrpAspThrGluIleGluThrGluLysProThrTyrValProAsnPhe--- 20

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
                       ||||||   ||||||:::         :::          |||::::::
Db         21 ---SerThrLeuGluThrGluAsnThrAspAsnTyr------SerAlaTyrSerAsnAsn 37

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ::::::::::::                                  :::   |||
Db         38 AspIleAsnAsn------------------------------GlnAsnTyrAspSerGlu 47

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              |||:::||||||||||||   |||      ::::::|||          |||
Db         48 ArgSerPheGlyAsnArg---GlyGlyTyrArgSerGluArgSerArgProSerAsnPhe 66

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
                                           |||||||||         ||||||   |||
Db         67 ------------------------------AsnArgGly---------SerArgThrGlu 73

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------------- 351
                 |||         ||||||         :::::::::
Db         74 ArgGlyArgGlyArgGlyPheGlyThrAsnArgAsnAspAsnTyrSerSerGluArgAsp 93

Qy        352 ---------GAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
                       :::|||         :::   |||||||||      |||||||||||
Db         94 ValPheGlyAspAspGluArgAspGlnArgArgGlyPheProGlyArgGlyGlyTyrAsn 113

Qy        403 GATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGA 462
                       :::   :::         :::|||      |||||||||   ||||||
Db        114 GlyAsnGluAspGlyGlnLysProAsnAlaPheArg------GlyArgGlyGlyPheArg 131

Qy        463 ---------------------------GGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCA 495
                                         |||   |||||||||||         :::
Db        132 AsnGluAsnGluGlnArgArgGlyPheGlyGluArgGlyGlyPheArgSerGluAsnGly 151

Qy        496 AATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGA 555
                 :::   |||                        |||   ||||||:::
Db        152 GlnArgAsnPheAsp--------------------AsnArgGlyAspPheGlyAsnSer 164

Qy        556 ---------------AGACCAGTATTAAGTGGCACAGGT---------AATGGTGATACT 591
                             ||||||      |||   |||         |||   ||||||
Db        165 GlyGluGluGluAspArgProArgSerTyrGlyArgGlyGlyPheAsnAsnSerAspThr 184

Qy        592 TCT------------CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGT 639
                 :::               |||   ||||||:::   |||||||||||||||
Db        185 GlyGlyArgGlyArgArgGlyGlyArgGlyGlyGlySerGlnTyrGlyGlyTyrLysGly 204

Qy        640 TTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGA 699
                 ||||||||||||            |||      |||   ||||||:::   ||||||
Db        205 ArgAsnGluGluVal------------GlyValGluSerGlyLysSerGlnGluGluGly 220

Qy        700 GGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAG 759
                 |||      |||   :::   |||||||||||||||||||||||||||||:::
Db        221 AsnGlu---LysAspGluLysProLysLysValThrTyrIleProProProProProAsp 239

Qy        760 GATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACT 819
                 ||||||:::||||||      ||||||:::||||||||||||||||||||||||
Db        240 GlyGluAspAsnIlePheArgGlnTyrGlnSerGlyIleAsnPheAspLysTyrAspGlu 259

Qy        820 ATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCT 879
              ||||||||||:::|||:::|||      |||   ||||||||||||||||||||||||
Db        260 IleLeuValAspValThrGlyLysAspValProProAlaIleLeuThrPheGluGluAla 279

Qy        880 AATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCT 939
              |||||||||:::||||||      |||:::|||:::|||||||||   ||||||||||||
Db        280 AsnLeuCysGluThrLeuArgArgAsnValAlaArgAlaGlyTyrValLysLeuThrPro 299

Qy        940 GTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAA 999
              |||||||||:::||||||||||||||||||:::|||||||||||||||||||||||||||
Db        300 ValGlnLysHisSerIleProIleIleMetAlaGlyArgAspLeuMetAlaCysAlaGln 319

Qy       1000 ACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGAT 1059
              ||||||||||||||||||||||||||||||||||||||||||::::::||||||::::::
Db        320 ThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerTyrMetMetAsnGlu 339

Qy       1060 GGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCA 1119
              ||||||||||||||||::::::   :::||||||||||||||||   ||||||:::||||||
```

```
Db        340 GlyIleThrAlaSerGlnTyrLeuGlnLeuGlnGluProGluAlaIleIleIleAlaPro 359

Qy       1120 ACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGT 1179
              |||||||||||:::||||||||||||||:::|||||||||||||||:::||||||||
Db        360 ThrArgGluLeuIleAsnGlnIleTyrLeuAspAlaArgLysPheSerTyrGlyThrCys 379

Qy       1180 GTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTA 1239
              ||||||  ||||||:::|||||||||   |||        |||::::::|||    :::
Db        380 ValArgProValValValTyrGlyGlyIleGlnProValHisAlaMetArgAspValGlu 399

Qy       1240 CAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAA 1299
              :::||||||||||||||||||||||||||||||||||||:::|||||||:::    ||||||
Db        400 LysGlyCysAsnIleLeuCysAlaThrProGlyArgLeuLeuAspIleValSerLysGlu 419

Qy       1300 AAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGAT 1359
              |||||||||||    :::::::::||||||||||||||||||||||||||||||||||||
Db        420 LysIleGlyLeuSerLysLeuArgTyrLeuValLeuAspGluAlaAspArgMetLeuAsp 439

Qy       1360 ATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAA 1419
              ||||||||    ||||||::::::||||||::::::    |||||||||||:::||||||
Db        440 MetGlyPheAlaProGluIleGluLysLeuMetThrLysProGlyMetProThrLysGlu 459

Qy       1420 CAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCA 1479
              :::||||||||||||||||||||||||||||:::||||||||||||:::|||||||||:::
Db        460 LysArgGlnThrLeuMetPheSerAlaThrTyrProGluGluIleArgArgLeuAlaSer 479

Qy       1480 GAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGA 1539
              :::||||||||||    :::||||||||||    ||||||   |||||||||||||||
Db        480 AsnTyrLeuLysSerGluHisLeuPheValValValGlyLeuValGlyGlyAlaCysSer 499

Qy       1540 GATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAA 1599
              ||||||  |||||||||||::::::   :::         |||  ||||||||:::|||
Db        500 AspValAlaGlnThrValLeuGluMetArgGluAsnGlyLysMetGluLysLeuLeuGlu 519

Qy       1600 ATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCA 1659
              ||||||::::::            |||||||||||:::||||||  ||||||||||||||
Db        520 IleLeuLysSerSerGluLysGluArgThrMetIlePheValAsnThrLysLysLysAla 539

Qy       1660 GATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGAT 1719
              ||||||   |||   :::||||||||||||||    |||:::||||||||||||||||||
Db        540 AspPheIleAlaGlyTyrLeuCysGlnGluLysPheSerSerThrSerIleHisGlyAsp 559

Qy       1720 CGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTT 1779
              ||||||||    :::||||||   ||||||  |||||||||   |||||||||    |||
Db        560 ArgGluGlnTyrGlnArgGluSerAlaLeuTrpAspPheArgThrGlyLysCysThrVal 579

Qy       1780 CTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATC 1839
              :::|||   |||:::||||||||||||||||||||||||||||||||||||||||||||
Db        580 IleValCysThrAlaValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIle 599

Qy       1840 AATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
              |||:::|||:::|||      :::|||||||||||||||||||||||||||||||||||
Db        600 AsnTyrAspValProLysGluValAspGluTyrValHisArgIleGlyArgThrGlyArg 619

Qy       1900 TGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATTTAGCA 1959
              |||||||||||||||:::|||  ||||||||::::::::   |||:::    :::|||
Db        620 CysGlyAsnThrGlyLysAlaThrSerPhePheAsnValGlnAspAspHisValIleAla 639

Qy       1960 CAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAA 2019
              :::||||||||||||:::||||||||||||    |||:::||||||||||||||||||||
Db        640 ArgProLeuValLysIleLeuThrAspAlaHisGlnGluValProAlaTrpLeuGluGlu 659

Qy       2020 ATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCA 2079
              ||||||                    |||   |||          |||     |||
Db        660 IleAla--------------------PheGlyGlyHisGlyAlaLeuAsnSerPheTyr 672

Qy       2080 TCAGTTGATACC---------AGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCT 2130
              :::    |||:::         :::   |||   :::   :::   |||      |||:::
Db        673 AlaAlaAspSerMetGlyGluGlnAlaGlyGlyAsnAlaValThrThrProSerPheAla 692

Qy       2131 TCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
                                          ::::::  |||||||||
Db        693 GlnGlu------------------GluGluAlaSerTrpAsp 700
```

```
RESULT 6
ABB59954
ID   ABB59954 standard; Protein; 661 AA.
XX
AC   ABB59954;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 6654.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL04057.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 6654; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   661 AA;

Alignment Scores:
Pred. No.:              5.07e-114      Length:        661
Score:                  1290.00        Matches:       314
Percent Similarity:     56.46%         Conservative:  101
Best Local Similarity:  42.72%         Mismatches:    230
Query Match:            32.63%         Indels:        90
DB:                     22             Gaps:          24

US-09-714-865-15 (1-2172) x ABB59954 (1-661)

Qy         10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG--- 66
                 :::||||||:::  |||                           ||||||   :::
Db          3 AspAspTrpAspAspGlu---------------------------ProIleValAspThr 13

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              :::             |||:::      |||:::   :::          :::   |||
Db         14 ArgGlyAlaArgGlyGlyAspTrpSerAspAspGluAspThrAlaLysSerPheSerGly 33

Qy        127 GAAATGGAT---GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCT 183
              |||   :::   ||||||                                 |||
Db         34 GluAlaGluGlyAspGlyValGlyGly---------------SerGlyGlyGluGlyGly 48
```

```
Qy        184 GGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACA 243
              |||          |||||||||||
Db         49 GlyTyrGlnGlyGlyAsnArgAsp------------------------------Val 57

Qy        244 ATGGGTGGTTTTGGAGTTGGAAAGAGTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTT 303
                 |||      |||   |||:::         |||      |||:::
Db         58 PheGlyArgIleGlyGlyGlyArgGlyGlyGlyAlaGlyGlyTyrArgGlyGlyAsnArg 77

Qy        304 GAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCA 363
              :::    |||         |||      ||||||       |||
Db         78 AspGlyGlyGlyPheHisGlyGlyArgArgGluGlyGluArgAspPhe----------- 93

Qy        364 ACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCT 423
                            ||||||         ||||||:::|||    |||
Db         94 ---------ArgGly------GlyGluGlyGlyPheArgGlyGlyGlnGlyGlySerArg 108

Qy        424 TCAGGGCCA---TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
                 |||         ||||||||:::|||   |||||||||     |||||||||
Db        109 GlyGlyGlnGlyGlySerArgGlyGlyGlnGlyGlyPheArgGlyGlyGluGlyGlyPhe 128

Qy        481 ---GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGT 537
                 |||         |||   |||   |||                :::|||    |||
Db        129 ArgGlyArgLeuTyrGluAsnGluAspGlyAsp--------------GluArgArgGly 143

Qy        538 GGCCTT--------------TTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAAT 582
                 |||                     |||   ||||||     |||
Db        144 ArgLeuAspArgGluGluArgGlyGlyGluArgArgGlyArgLeuAsp----------- 159

Qy        583 GGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGA------GGTGGTTAC--- 633
                          :::   |||   |||         ||||||      ||||||:::
Db        160 ---------ArgGluGluArgGlyGlyGluArgGlyGluArgGlyAspGlyGlyPheAla 176

Qy        634 ---AAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAA 690
                 :::         ||||||:::
Db        177 ArgArgArgArgAsnGluAspAspIleAsnAsnAsn------------------------ 188

Qy        691 GCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCT 750
                                |||    :::       :::      |||||||||||
Db        189 ------AsnAsnIleValGluAspValGluArgLysArgGluPheTyrIleProProGlu 206

Qy        751 CCACCTGAGGATGAGGACTCCATCTTTGCACAT---TATCAGACAGGCATAAACTTCGAC 807
              |||      |||         ||||||:::             :::||||||:::|||
Db        207 ProSerAsnAspAlaIleGluIlePheSerSerGlyIleAlaSerGlyIleHisPheSer 226

Qy        808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
              ||||||:::    |||    |||:::|||:::|||    |||    |||       |||
Db        227 LysTyrAsnAsnIleProValLysValThrGlySerAspValProGlnProIleGlnHis 246

Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
              |||         |||:::|||          :::    :::|||:::    |||:::||||||
Db        247 PheThrSerAlaAspLeuArgAspIleIleIleAspAsnValAsnLysSerGlyTyrLys 266

Qy        928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
                       ||||||:::||||||    |||||||||:::|||    :::|||||||||||||||
Db        267 IleProThrProIleGlnLysCysSerIleProValIleSerSerGlyArgAspLeuMet 286

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        287 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerLys 306

Qy       1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
              ::::::    |||                        ||||||     |||:::    :::
Db        307 LeuLeuGluAsp-----------ProHisGluLeuGluLeuGlyArgProGlnValVal 322

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
              ||||||:::||||||||||||||||         ||||||:::    |||||||||||||||:::
Db        323 IleValSerProThrArgGluLeuAlaIleGlnIlePheAsnGluAlaArgLysPheAla 342

Qy       1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
              |||    :::    ::::::        ::::::||||||||||||          |||
Db        343 PheGluSerTyrLeuLysIleGlyIleValTyrGlyGlyThrSerPheArgHisGlnAsn 362

Qy       1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
                    |||   :::||||||:::::::::    ||||||||||||||||||:::|||
Db        363 GluCysIleThrArgGlyCysHisValValIleAlaThrProGlyArgLeuLeuAspPhe 382

Qy       1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
              :::   :::         |||      :::    ::::::::::||||||||||||||||||
Db        383 ValAspArgThrPheIleThrPheGluAspThrArgPheValValLeuAspGluAlaAsp 402

Qy       1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
              ||||||||||||||||||          :::|||:::::::::::::::          |||
Db        403 ArgMetLeuAspMetGlyPheSerGluAspMetArgArgIleMetThrHisValThrMet 422

Qy       1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
                    :::   :::   ||||||||||||||||||||||||||||||||||||||||||
Db        423 ------ArgProGluHisGlnThrLeuMetPheSerAlaThrPheProGluGluIleGln 440

Qy       1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGT 1527
              |||:::|||   ||||||||||||    ||||||:::||||||||||:::|||    ||||||
Db        441 ArgMetAlaGlyGluPheLeuLys---AsnTyrValPheValAlaIleGlyIleValGly 459

Qy       1528 GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAA 1587
              |||||||||   ||||||:::||||||:::      :::|||    :::::::::||||||
Db        460 GlyAlaCysSerAspValLysGlnThrIleTyrGluValAsnLysTyrAlaLysArgSer 479

Qy       1588 AAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACT 1647
              ||||||:::|||||||||               |||      |||:::|||||||||||||||
Db        480 LysLeuIleGluIleLeuSerGluGlnAlaAsp---GlyThrIleValPheValGluThr 498

Qy       1648 AAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGT 1707
              |||:::    |||||||||    |||:::||||||    :::::::::       |||||||||
Db        499 LysArgGlyAlaAspPheLeuAlaSerPheLeuSerGluLysGluPheProThrThrSer 518

Qy       1708 ATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGA 1767
              |||||||||||||||   |||   :::||||||||||||||||||    ||||||:::   |||
Db        519 IleHisGlyAspArgLeuGlnSerGlnArgGluGlnAlaLeuArgAspPheLysAsnGly 538

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
                    ||||||:::||||||||||||||||||:::||||||||||||||||||:::|||:::
Db        539 SerMetLysValLeuIleAlaThrSerValAlaSerArgGlyLeuAspIleLysAsnIle 558

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              :::|||||||||||||:::|||:::||||||     ||||||:::||||||||||||||||||
Db        559 LysHisValIleAsnTyrAspMetProSerLysIleAspAspTyrValHisArgIleGly 578

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||    ||||||   |||||||||   ||||||||||||    |||    |||
Db        579 ArgThrGlyArgValGlyAsnAsnGlyArgAlaThrSerPhePheAspProGluLysAsp 598

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
                    :::|||      |||||||||:::|||        :::    |||    ||||||
Db        599 ArgAlaIleAlaAlaAspLeuValLysIleLeuGluGlySerGlyGlnThrValProAsp 618

Qy       2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
              :::|||                        |||         |||     |||
Db        619 PheLeuArg--------------ThrCysGlyAlaGlyGlyAspGlyGlyTyrSerAsn 633

Qy       2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
                    |||      ||||||   |||   |||:::            |||      |||
Db        634 GlnAsnPheGlyGlyValAspValArg---GlyArgGly------AsnTyrValGly--- 649

Qy       2128 TCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
                          |||   |||         ::::::|||   ||||||
Db        650 ---------AspAlaThrAsnValGluGluGluGluGlnTrpAsp 661
```

RESULT 7

```
AAE02422
ID   AAE02422 standard; Protein; 661 AA.
XX
AC   AAE02422;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Fruit fly vasa protein.
```

```
XX
KW   Vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial; fruit fly.
XX
OS   Drosophila melanogaster.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;    99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 61-62; 66pp; English.
XX
CC   The present sequence is fruit fly vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   661 AA;

Alignment Scores:
Pred. No.:              1.71e-112     Length:        661
Score:                  1274.00       Matches:       314
Percent Similarity:     56.73%        Conservative:  103
Best Local Similarity:  42.72%        Mismatches:    228
Query Match:            32.22%        Indels:        90
DB:                     22            Gaps:          25

US-09-714-865-15 (1-2172) x AAE02422 (1-661)

Qy         10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG--- 66
              :::||||||:::   |||                           ||||||    :::
Db          3 AspAspTrpAspAspGlu---------------------------ProIleValAspThr 13

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              :::              |||:::     |||:::   :::         :::   |||
Db         14 ArgGlyAlaArgGlyGlyAspTrpSerAspAspGluAspThrAlaLysSerPheSerGly 33

Qy        127 GAAATGGAT---GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCT 183
              |||   :::   ||||||                             |||
Db         34 GluAlaGluGlyAspGlyValGlyGly---------------SerGlyGlyGluGlyGly 48

Qy        184 GGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACA 243
              |||            ||||||||||||
Db         49 GlyTyrGlnGlyGlyAsnArgAsp------------------------------Val 57

Qy        244 ATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTT 303
              |||        |||   |||:::         |||        |||:::
Db         58 PheGlyArgIleGlyGlyGlyArgGlyGlyGlyAlaGlyGlyTyrArgGlyGlyAsnArg 77

Qy        304 GAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCA 363
              :::   |||         |||      ||||||         |||
Db         78 AspGlyGlyGlyPheHisGlyGlyArgArgGluGlyGluArgAspPhe------------ 93

Qy        364 ACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCT 423
                       ||||||            ||||||:::|||   |||
Db         94 ---------ArgGly------GlyGluGlyGlyPheArgGlyGlyGlnGlyGlySerArg 108

Qy        424 TCAGGGCCA---TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
                 |||         |||||||||:::|||   |||||||||         |||||||||
Db        109 GlyGlyGlnGlyGlySerArgGlyGlyGlnGlyGlyPheArgGlyGlyGlyGluGlyGlyPhe 128

Qy        481 ---GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGT 537
                 |||         |||   |||   |||                :::|||   |||
Db        129 ArgGlyArgLeuTyrGluAsnGluAspGlyAsp--------------GluArgArgGly 143

Qy        538 GGCCTT--------------TTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAAT 582
                 |||                |||   ||||||      |||
Db        144 ArgLeuAspArgGluGluArgGlyGlyGluArgArgGlyArgLeuAsp----------- 159

Qy        583 GGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGA------GGTGGTTAC--- 633
                       :::   |||   |||         ||||||      ||||||:::
Db        160 ---------ArgGluGluArgGlyGlyGluArgGlyGluArgGlyAspGlyGlyPheAla 176

Qy        634 ---AAAGGTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAA 690
                 :::      |||||||:::                |||:::      :::
Db        177 ArgArgArgArgAsnGluAspAspIle---------------AsnAsnAsnAsnAsnIle 191

Qy        691 GCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCT 750
              ||||||               |||   :::      :::      ||||||||||||
Db        192 AlaGlu---------------AspValGluArgLysArgGluPheTyrIleProProGlu 206

Qy        751 CCACCTGAGGATGAGGACTCCATCTTTGCACAT---TATCAGACAGGCATAAACTTCGAC 807
              |||      |||         ||||||:::            :::||||||:::|||
Db        207 ProSerAsnAspAlaIleGluIlePheSerSerGlyIleAlaSerGlyIleHisPheSer 226

Qy        808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
              ||||||:::   |||   |||:::|||:::|||   |||   |||      |||
Db        227 LysTyrAsnAsnIleProValLysValThrGlySerAspValProGlnProIleGlnHis 246

Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
              |||      |||:::|||      :::   :::|||:::   |||:::|||:::
Db        247 PheThrSerAlaAspLeuArgAspIleIleIleAspAsnValAsnLysSerGlyPheLys 266

Qy        928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
                    ||||||:::||||||   |||||||||:::|||   :::|||||||||||||||
Db        267 IleProThrProIleGlnLysCysSerIleProValIleSerSerGlyArgAspLeuMet 286

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        287 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerLys 306

Qy       1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
              ::::::   |||                           ||||||      |||:::   :::
Db        307 LeuLeuGluAsp------------ProHisGluLeuGluLeuGlyArgProGlnValVal 322

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
              ||||||:::|||||||||||||||         ||||||:::   |||||||||||||||:::
Db        323 IleValSerProThrArgGluLeuAlaIleGlnIlePheAsnGluAlaArgLysPheAla 342

Qy       1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
              |||   :::   ::::::      ::::::||||||||||||         |||
Db        343 PheGluSerTyrLeuLysIleGlyIleValTyrGlyGlyThrSerPheArgHisGlnAsn 362
```

```
Qy       1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
                     |||   :::||||||:::::::::   ||||||||||||||||||:::|||
Db        363 GluCysIleThrArgGlyCysHisValValIleAlaThrProGlyArgLeuLeuAspPhe 382

Qy       1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
                :::   :::      |||      :::       ::::::::||||||||||||||||||
Db        383 ValAspArgThrPheIleThrPheGluAspThrArgPheValValLeuAspGluAlaAsp 402

Qy       1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
              ||||||||||||||||||||            :::|||:::::::::::::::        |||
Db        403 ArgMetLeuAspMetGlyPheSerGluAspMetArgArgIleMetThrHisValThrMet 422

Qy       1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
                     :::   :::   ||||||||||||||||||||||||||||||||||||||||
Db        423 ------ArgProGluHisGlnThrLeuMetPheSerAlaThrPheProGluGluIleGln 440

Qy       1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGT 1527
              |||:::|||   ||||||||||||   ||||||:::   ||||||:::|||    ||||||
Db        441 ArgMetAlaGlyGluPheLeuLys---AsnTyrValSerValAlaIleGlyIleValGly 459

Qy       1528 GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAA 1587
              |||||||||   ||||||:::|||||||:::   :::|||   :::::::::||||||
Db        460 GlyAlaCysSerAspValLysGlnThrIleTyrGluValAsnLysTyrAlaLysArgSer 479

Qy       1588 AAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACT 1647
              ||||||:::|||||||||               |||         |||:::|||||||||||||||
Db        480 LysLeuIleGluIleLeuSerGluGlnAlaAsp---GlyThrIleValPheValGluThr 498

Qy       1648 AAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGT 1707
              |||:::   |||||||||   |||:::||||||   :::::::::      |||||||||
Db        499 LysArgGlyAlaAspPheLeuAlaSerPheLeuSerGluLysGluPheProThrThrSer 518

Qy       1708 ATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGA 1767
              |||||||||||||||   |||   :::|||||||||||||||   ||||||:::   |||
Db        519 IleHisGlyAspArgLeuGlnSerGlnArgGluGlnAlaLeuArgAspPheLysAsnGly 538

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
                     ||||||:::||||||||||||||||:::||||||||||||||||:::|||:::
Db        539 SerMetLysValLeuIleAlaThrSerValAlaSerArgGlyLeuAspIleLysAsnIle 558

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              :::||||||||||||:::|||:::||||||   ||||||:::||||||||||||||||||
Db        559 LysHisValIleAsnTyrAspMetProSerLysIleAspAspTyrValHisArgIleGly 578

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              |||||||||      ||||||   |||||||||   ||||||||||||   |||   |||
Db        579 ArgThrGlyCysValGlyAsnAsnGlyArgAlaThrSerPhePheAspProGluLysAsp 598

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
                     :::|||      |||||||||:::|||      :::   |||   ||||||
Db        599 ArgAlaIleAlaAlaAspLeuValLysIleLeuGluGlySerGlyGlnThrValProAsp 618

Qy       2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
              :::|||                           |||      |||      |||
Db        619 PheLeuArg---------------ThrCysGlyAlaGlyGlyAspGlyGlyTyrSerAsn 633

Qy       2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
                     |||      ||||||   |||   |||:::               |||      |||
Db        634 GlnAsnPheGlyGlyValAspValArg---GlyArgGly------AsnTyrValGly--- 649

Qy       2128 TCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
                             |||   |||      ::::::|||   ||||||
Db        650 ---------AspAlaThrAsnValGluGluGluGluGlnTrpAsp 661
```

```
RESULT 8
ABB64631
ID   ABB64631 standard; Protein; 798 AA.
XX
AC   ABB64631;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 20685.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL08734.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 20685; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   798 AA;

Alignment Scores:
Pred. No.:               1.94e-109     Length:         798
Score:                   1242.50       Matches:        313
Percent Similarity:      54.78%        Conservative:   111
Best Local Similarity:   40.44%        Mismatches:     231
Query Match:             31.42%        Indels:         119
DB:                      22            Gaps:           27

US-09-714-865-15 (1-2172) x ABB64631 (1-798)

Qy          4 GGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCC--------------- 48
              |||      |||:::         |||         |||:::
Db         24 GlySerAlaAspTyrSerGlyProIleThrSerLysThrSerThrAsnSerValThrGly 43

Qy         49 ------TATGTTCCCATATTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTT 102
                    ||||||||                  |||      |||:::|||
Db         44 GlyValTyrValPro------------ProHisLeuArgGlyGlyGlyGlyAsnAsn--- 58

Qy        103 AACAGGACTCCAGCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTC 162
              |||            |||   |||            |||
Db         59 AsnAlaAlaAspAlaGluSerGlnGlyGlnGlyGlnGly--------------------- 71

Qy        163 ATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAAT 222
                                |||:::   |||   :::|||
Db         72 -----------------GlnGlyGlnGlyPheAspSerArgSerGlyAsnProArgGln 85

Qy        223 AAG---CGAGATAATACATCCACAATGGGTGGTTTTGGAGTT---------GGAAAGAGT 270
              :::   ||||||      :::   ||||||   |||               |||
Db         86 GluThrArgAspProGlnGlnSerArgGlyGlyGlyGlyGluTyrArgArgGlyGlyGly 105

Qy        271 TTTGGAAACAGAGGTTTTTCA------AACAGCAGGTTTGAAGATGGTGATAGCTCTGGT 324
```

```
Db        106 GlyGlyGlyArgGlyPheAsnArgGlnSerGlyAspTyrGlyTyrGlySerGlyGlyGly 125

Qy        325 TTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCC 384
Db        126 GlyArgArgGlyGlyGlyGlyArgPheGluAspAsnTyrAsnGlyGlyGluPheAspSer 145

Qy        385 AAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGT 444
Db        146 ArgArgGlyGly------AspTrpAsnArgSerGlyGlyGlyGlyGlyGlyGlyArgGly 163

Qy        445 ---GGAAGAGGT---AGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGT------ 492
Db        164 PheGlyArgGlyProSerTyrArgGly------GlyGlyGlyGlySerGlySerAsnLeu 181

Qy        493 ---------------------------------------------CCAAATAATGAC--- 504
Db        182 AsnGluGlnThrAlaGluAspGlyGlnAlaGlnGlnGlnGlnGlnProArgAsnAspArg 201

Qy        505 ---TTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCA 561
Db        202 TrpGlnGluPro------------GluArgProAlaGlyPheAspGlySerGluGlyGly 217

Qy        562 GTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAA 621
Db        218 GlnSerAlaGly---------GlyAsnArgSerTyrAsnAsnArgGly---------Glu 231

Qy        622 CGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGG 681
Db        232 ArgGlyGly------------------------------GlyGlyTyrAsnSerArgTrp 241

Qy        682 AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATA 741
Db        242 Lys---------GluGlyGly------------------GlySerAsnValAspTyrThr 252

Qy        742 CCCCCTCCTCCACCTGAGGAT------GAGGACTCCATCTTTGCACATTATCAGACAGGC 795
Db        253 LysLeuGlyAlaArgAspGluArgLeuGluValGluLeuPheGlyValGlyAsnThrGly 272

Qy        796 ATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCA 855
Db        273 IleAsnPheAspLysTyrGluAspIleProValGluAlaThrGlyGlnAsnValProPro 292

Qy        856 GCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAA 915
Db        293 AsnIleThrSerPheAspAspValGlnLeuThrGluIleIleArgAsnAsnValAlaLeu 312

Qy        916 GCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGA 975
Db        313 AlaArgTyrAspLysProThrProValGlnLysHisAlaIleProIleIleIleAsnGly 332

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
Db        333 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuValPro 352

Qy       1036 ATTTTGGCTCATATGATGCATGATGGA-------------------ATAACTGCCAGTCGT 1077
Db        353 IleLeuAsnGlnMetTyrGluLeuGlyHisValProProProGlnSerThrArgGlnTyr 372

Qy       1078 TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAAC 1137
Db        373 SerArgArgLysGlnTyrProLeuGlyLeuValLeuAlaProThrArgGluLeuAlaThr 392

Qy       1138 CAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATA 1197
Db        393 GlnIlePheGluGluAlaLysLysPheAlaTyrArgSerArgMetArgProAlaValLeu 412

Qy       1198 TATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTA 1257
Db        413 TyrGlyGlyAsnAsnThrSerGluGlnMetArgGluLeuAspArgGlyCysHisLeuIle 432

Qy       1258 TGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAG 1317
Db        433 ValAlaThrProGlyArgLeuGluAspMetIleThrArgGlyLysValGlyLeuGluAsn 452

Qy       1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAA 1377
Db        453 IleArgPheLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGln 472

Qy       1378 ATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATG 1437
Db        473 IleArgArgIleValGluGlnLeuAsnMetProProThrGlyGlnArgGlnThrLeuMet 492

Qy       1438 TTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAAT 1497
Db        493 PheSerAlaThrPheProLysGlnIleGlnGluLeuAlaSerAspPheLeu---SerAsn 511

Qy       1498 TATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT 1557
Db        512 TyrIlePheLeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnThrIle 531

Qy       1558 CTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATA--- 1614
Db        532 LeuTrpValTyrGluProAspLysArgSerTyrLeuLeuAspLeuLeuSerSerIleArg 551

Qy       1615 ------------------GGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAA 1656
Db        552 AspGlyProGluTyrThrLysAspSerLeuThrLeuIlePheValGluThrLysLysGly 571

Qy       1657 GCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGT 1716
Db        572 AlaAspSerLeuGluGluPheLeuTyrGlnCysAsnHisProValThrSerIleHisGly 591

Qy       1717 GATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCA 1776
Db        592 AspArgThrGlnLysGluArgGluGluAlaLeuArgCysPheArgSerGlyAspCysPro 611

Qy       1777 GTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTT 1836
Db        612 IleLeuValAlaThrAlaValAlaAlaArgGlyLeuAspIleProHisValLysHisVal 631

Qy       1837 ATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGT 1896
Db        632 IleAsnPheAspLeuProSerAspValGluGluTyrValHisArgIleGlyArgThrGly 651

Qy       1897 CGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTA 1956
Db        652 ArgMetGlyAsnLeuGlyValAlaThrSerPhePheAsn---GluLysAsnArgAsnIle 670

Qy       1957 GCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAA 2016
Db        671 CysSerAspLeuLeuGluLeuIleGluThrLysGlnGluIleProSerPheMetGlu 690

Qy       2017 GAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGA------------ 2064
Db        691 AspMetSerSerAspArg---------GlyHisGlyGlyAlaLysArgAlaGlyArgGly 707

Qy       2065 ---------------GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACT 2109
Db        708 GlyGlyGlyArgTyrGlyGlyGlyPheGlySerArgAspTyrArgGln------------ 723

Qy       2110 TTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCA 2151
Db        724 ---SerSerGlyGlyGlyGlyGlyGlyArgSerGlyProPro 736
```

```
RESULT 9
ABB57045
ID   ABB57045 standard; Protein; 662 AA.
XX
AC   ABB57045;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Mouse ischaemic condition related protein sequence SEQ ID NO:68.
XX
KW   Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
```

```
KW   vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS   Mus musculus.
XX
PN   WO200188188-A2.
XX
PD   22-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-JP04192.
XX
PR   18-MAY-2000; 2000JP-0145977.
XX
PA   (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI   Ishikawa K,  Asai S,  Takahashi Y,  Nagata T,  Ishii Y;
XX
DR   WPI; 2002-034733/04.
DR   N-PSDB; ABI99243.
XX
PT   Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT   expression levels of particular genes defined in the specification or
PT   by determining the expression profile of a gene group comprising these
PT   genes -
XX
PS   Claim 2; Page 215-218; 2690pp; English.
XX
CC   The present invention describes a method for examining ischaemic
CC   conditions, comprising measuring the expression levels of particular
CC   genes (I) in a test sample or determining the expression profile of a
CC   gene group in the sample comprising genes selected from (I). The method
CC   is useful for examining the ischaemic condition (e.g. compressive
CC   ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC   expression levels of particular genes (ABI99202 to ABI99912, encoding
CC   the protein sequences in ABB57020 to ABB57374) or by determining the
CC   expression profile of a gene group comprising these genes. The
CC   expression levels or expression profiles produced by these genes are
CC   used as an indicator when screening for ischaemic condition-improving
CC   drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC   represent PCR primers for a mouse ischaemic condition related sequence,
CC   which are used in the exemplification of the present invention.
XX
SQ   Sequence   662 AA;

Alignment Scores:
Pred. No.:             3.44e-108      Length:        662
Score:                 1229.00        Matches:       297
Percent Similarity:    57.59%         Conservative:  90
Best Local Similarity: 44.20%         Mismatches:    183
Query Match:           31.08%         Indels:        102
DB:                    23             Gaps:          22

US-09-714-865-15 (1-2172) x ABB57045 (1-662)

Qy        214 GAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTGGAGTTGGAAAGAGT--- 270
              :::  |||    ||||||  |||     ||||||              |||
Db         20 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 37

Qy        271 ---------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
                             ||||||      :::        |||   |||   |||
Db         38 TyrIleProProHisLeuArgAsnArgGlu---AlaThrLysGlyPheTyrAspLysAsp 56

Qy        316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGA 375
              |||||||||:::                 |||      :::
Db         57 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSer--------------- 71

Qy        376 GGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATAC 435
                 |||   ||||||   |||          |||          |||
Db         72 ---PheGlySerArgGlyAspSerArg------GlyLysSerSerPhePheGlyAspArg 88

Qy        436 AGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGCCGTGGAGGATTT---GGTCTA 486
                 |||  ||||||  |||        ||||||  ||||||   :::   |||:::
Db         89 GlySerGlySerArgGlyArgPheAspAspArgGly---ArgGlyAspTyrAspGlyIle 107

Qy        487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 546
              |||
Db        108 Gly--------------------------------------------------------- 108

Qy        547 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 606
                                            |||  ||||||            ||||||
Db        109 -----------------------------GlyArgGlyAsp------------ArgSer 114

Qy        607 GGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGC 663
              |||  |||     |||||||||||            |||
Db        115 GlyPheGlyLysPheGluArgGlyGly-----------AsnSerArgTrpCysAspLys 130

Qy        664 TCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGA 723
              |||   ::::::   |||
Db        131 SerAspGluAspAspTrpSerLys------------------------------------ 138

Qy        724 CCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT---GAGGACTCCATCTTTGCA 780
                                |||  ||||||  |||       |||       :::|||:::
Db        139 -----------------ProLeuProProSerGluArgLeuGluGlnGluLeuPheSer 152

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
                 ||||||||||||||:::|||||||  |||  ||||||    :::|||
Db        153 GlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIleProValGluAlaThrGly 172

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
              ::::::   ||||||  |||   :::|||   :::       :::   :::   :::
Db        173 AsnAsnCysProProHisIleGluSerPheSerAspValGluMetGlyGluIleIleMet 192

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
                 ||||||            ||||||:::   |||||||||||||::::::||||||
Db        193 GlyAsnIleGluLeuThrArgTyrThrArgProThrProValGlnLysHisAlaIlePro 212

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              ||||||         |||||||||||||||||||||||||||||||||||||||||||
Db        213 IleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 232

Qy       1021 GCTTTTCTCCCTACCAATTTTGGCTCATATGATGCATGATGGA-----------ATAACT 1068
              ||||||||||||||||||||||:::   :::         ||||||           :::
Db        233 AlaPheLeuLeuProIleLeuSerGlnIleTyrAlaAspGlyProGlyGluAlaLeuArg 252

Qy       1069 GCCAGTCGT---------------TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTA 1113
              |||   :::               :::      |||   |||       :::::::::
Db        253 AlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyrProIleSerLeuValLeu 272

Qy       1114 GCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTTCTTTTGGG 1173
              |||||||||||||||||         |||||||||   |||||||||||||||||:::
Db        273 AlaProThrArgGluLeuAlaValGlnIleTyrGluGluAlaArgLysPheSerTyrArg 292

Qy       1174 ACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGGGACATTCAATTCGACAA 1233
              :::   ||||||      |||:::|||||||||   ::::::|||       ||||||
Db        293 SerArgValArgProCysValValTyrGlyGlyAlaGluIleGlyGlnGlnIleArgAsp 312

Qy       1234 ATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGC 1293
              :::   :::||||||::::::|||   ||||||||||||||||||:::|||::::::
Db        313 LeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArgLeuValAspMetMetGlu 332

Qy       1294 AAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATG 1353
              :::   ||||||||||||            ||||||||||||||||||||||||||||||
Db        333 ArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeuAspGluAlaAspArgMet 352

Qy       1354 TTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCA 1413
              ||||||||||||||   |||::::::::::::::::::            ||||||
Db        353 LeuAspMetGlyPheGluProGlnIleArgArgIleValGluGlnAspThrMetProPro 372

Qy       1414 AAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTG 1473
              |||       |||   |||:::||||||||||||||||||||:::|||||||||   |||
Db        373 LysGlyValArgHisThrMetMetPheSerAlaThrPheProLysGluIleGlnMetLeu 392

Qy       1474 GCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCA 1533
              |||   :::||||||         |||:::|||:::|||||||||:::||||||
Db        393 AlaArgAspPheLeu---AspGluTyrIlePheLeuAlaValGlyArgValGlySerThr 411

Qy       1534 TGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTC 1593
```

```
              ::::::  |||   |||:::   |||   :::      ||||||      |||
Db        412 SerGluAsnIleThrGlnLysValValTrpValGluGluIleAspLysArgSerPheLeu 431

Qy       1594 GTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATGGTCTTTGTTGAAACTAAG 1650
              :::::::::|||        |||    |||       |||:::|||||||||||||||||||
Db        432 LeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeuValPheValGluThrLys 451

Qy       1651 AAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATC 1710
              |||   ||||||              ||||||     |||       :::   |||||||||
Db        452 LysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGlyTyrAlaCysThrSerIle 471

Qy       1711 CATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAG 1770
              ||||||||||||   ||||||:::||||||:::||||||          ||||||   ||||||
Db        472 HisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHisGlnPheArgSerGlyLys 491

Qy       1771 TGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAA 1830
                 |||:::||||||||||||:::||||||||||||||||||||||||   ||||||:::
Db        492 SerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeuAspIleSerAsnValLys 511

Qy       1831 CATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGT 1890
              |||||||||||||||||||||||||   .|||:::|||||||||||||||||||||||||
Db        512 HisValIleAsnPheAspLeuProSerAspIleGluGluTyrValHisArgIleGlyArg 531

Qy       1891 ACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAAC 1950
              |||||||||   ||.||||   |||   |||   ||||||||||:::   |||   :::
Db        532 ThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePheAsn---GluArgAsnIle 550

Qy       1951 CATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGG 2010
              ::::::   :::   |||:::   :::|||   :::|||:::|||:::||||||:::|||
Db        551 AsnIleThrLysAspLeuLeuAspLeuLeuValGluAlaLysGlnGluValProSerTrp 570

Qy       2011 TTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAAC 2070
              ||||||   :::||||||        :::            :::   ||||||:::||||||
Db        571 LeuGluAsnMetAlaPheGluHisHis---------TyrLysGlySerSerArgGlyArg 587

Qy       2071 GTG-----------------------TTTGCATCAGTTGATACCAGAAAGGGCAAGAGC 2106
                                        |||   :::   |||   |||:::      |||
Db        588 SerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAspTyrArgGln---SerSer 606

Qy       2107 ACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                    :::::::::   ||||||||||||||||||
Db        607 GlyAlaSerSerSerSerPheSerSerSerArgAla 618
```

```
RESULT 10
AAW81502
ID   AAW81502 standard; Protein; 662 AA.
XX
AC   AAW81502;
XX
DT   02-FEB-1999  (first entry)
XX
DE   Dead Box X (DBX) gene short transcript amino acid sequence.
XX
KW   Non-recombining region; human; Y chromosome; X homologue; testis; DBX;
KW   infertility; sperm; gene alteration; inhibitor; Dead Box X.
XX
OS   Homo sapiens.
XX
PN   WO9846747-A2.
XX
PD   22-OCT-1998.
XX
PF   10-APR-1998;   98WO-US07115.
XX
PR   11-APR-1997;   97US-0041877.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Lahn BT,  Page DC;
XX
DR   WPI; 1998-568729/48.
DR   N-PSDB; AAV69632.
XX
PT   Novel genes in the non-combining region of Y chromosome - useful to
PT   diagnose if male infertility or reduced sperm count has a genetic
PT   basis
XX
PS   Disclosure; Fig 3A-B; 54pp; English.
XX
CC   This represents the amino acid sequence of the Dead Box X (DBY) gene
CC   short transcript. The invention relates to genes occurring on the non-
CC   recombining region of the human Y chromosome. The sequences fall into two
CC   classes: (1) X-homologous DNA which are expressed in many organs, having
CC   functional X homologues and (2) testis-specific DNA sequences. Y
CC   chromosomal DNA from males with known conditions such as infertility and
CC   reduced sperm count can be assessed using the invention to determine
CC   whether the condition is associated with or caused by the occurrence of
CC   the gene or gene alteration. Candidate inhibitors of the enzymatic
CC   activity of the genes can be assessed using in vitro assays.
XX
SQ   Sequence   662 AA;

Alignment Scores:
Pred. No.:              7.48e-107     Length:         662
Score:                  1215.00       Matches:        293
Percent Similarity:     56.64%        Conservative:   91
Best Local Similarity:  43.22%        Mismatches:     180
Query Match:            30.73%        Indels:         114
DB:                     19            Gaps:           20

US-09-714-865-15 (1-2172) x AAW81502 (1-662)
```

```
Qy        214 GAGTGTAATAAGCGAGATAATACATCCACAATGGTGGTTTTGGAGTTGGAAAGAGT--- 270
              :::   |||         ||||||   |||         ||||||         |||
Db         20 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 37

Qy        271 --------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
                                  ||||||      :::            |||   |||   |||
Db         38 TyrIleProProHisLeuArgAsnArgGluAlaThrArgGly---PheTyrAspLysAsp 56

Qy        316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------------------GAA 354
              |||||||||:::                  |||
Db         57 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
              |||:::      ::::::      ||||||   ||||||
Db         77 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer------------------ 90

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGC 468
                                        |||   ||||||   |||      ||||||
Db         91 --------------------------GlySerArgGlyArgPheAspAspArgGlyArg 101

Qy        469 CGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG 528
                             |||:::||||||
Db        102 SerAspTyrAspGlyIleGlySer------------------------------------ 109

Qy        529 CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGAT 588
                                                                     ||||||
Db        110 ---------------------------------------------------ArgGlyAsp 112

Qy        589 ACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAAT 645
                          |||||||||   |||      |||||||||||||            |||
Db        113 ------------ArgSerGlyPheGlyLysPheGluArgGlyGly-----------Asn 124

Qy        646 GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
                                  |||.   ::::::   |||
Db        125 SerArgTrpCysAspLysSerAspGluAspAspTrpSerLys------------------ 138

Qy        706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--- 762
                                                        |||   ||||||   |||
Db        139 --------------------------------------ProLeuProProSerGluArgLeu 146

Qy        763 GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATT 822
              |||      :::|||:::         |||||||||||||||:::|||||||||   |||
Db        147 GluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIle 166
```

```
Qy        823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                 |||||    :::|||::::::   |||||   |||    :::|||    :::
Db        167 ProValGluAlaThrGlyAsnAsnCysProProHisIleGluSerPheSerAspValGlu 186

Qy        883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
              :::   :::   :::         |||||            ||||||:::   ||||||||
Db        187 MetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgProThrProVal 206

Qy        943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
              ||||||::::::||||||||||         ||||||||||||||||||||||||
Db        207 GlnLysHisAlaIleProIleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThr 226

Qy       1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTGGCTCATATGATGCATGATGGA 1062
              ||||||||||||||||||||||||||||||||||||:::    :::      ||||||
Db        227 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 246

Qy       1063 ------------ATAACTGCCAGTCGT---------------TTTAAAGAGTTGCAGGAA 1095
                          :::  |||  :::                   :::      |||
Db        247 ProGlyGluAlaLeuArgAlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyr 266

Qy       1096 CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTGGAAGCC 1155
              |||      :::::::::||||||||||||||||||        ||||||||   ||||||
Db        267 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluGluAla 286

Qy       1156 AGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTG 1215
              |||||||||||:::    :::    ||||||     |||:::|||||||||       :::
Db        287 ArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGlyAlaAspIle 306

Qy       1216 GGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGA 1275
              |||     ||||||   :::   :::||||||::::::|||   ||||||||||||||
Db        307 GlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArg 326

Qy       1276 CTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTG 1335
              |||:::|||::::::   :::   ||||||||||||         ||||||||||||||
Db        327 LeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeu 346

Qy       1336 GATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT 1395
              ||||||||||||||||||||||||||||||||||     |||:::::::::::::::::
Db        347 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArgIleValGlu 366

Qy       1396 TGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCA 1455
                        ||||||  |||      |||   |||:::|||||||||||||||||||||
Db        367 GlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAlaThrPhePro 386

Qy       1456 GAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT 1515
              :::|||||||||   ||||||    :::||||||           |||:::|||:::||||||
Db        387 LysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePheLeuAlaVal 405

Qy       1516 GGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTC 1575
                 |||:::||||||              ::::::   |||   |||:::   |||    :::
Db        406 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpValGluGluSer 425

Qy       1576 TCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATG 1632
                 ||||||   |||:::::::::|||            |||   |||      |||:::
Db        426 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeu 445

Qy       1633 GTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAA 1692
              |||||||||||||||||||||   ||||||        ||||||        |||
Db        446 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGly 465

Qy       1693 ATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGA 1752
                 :::   ||||||||||||||||||||   ||||||:::||||||:::||||||
Db        466 TyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHis 485

Qy       1753 GATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTG 1812
                 ||||||   ||||||   |||:::||||||||||||:::||||||||||||||||||
Db        486 GlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 505

Qy       1813 GATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATAT 1872
              ||||||   ||||||:::|||||||||||||||||||||||||||   |||:::||||||
Db        506 AspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIleGluGluTyr 525

Qy       1873 GTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTT 1932
              ||||||||||||||||||||||||||||||   ||||||   |||   |||   ||||||||
Db        526 ValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 545

Qy       1933 GATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAA 1992
              :::   |||   :::   ::::::   :::   |||:::   :::|||   :::|||:::
Db        546 Asn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuValGluAlaLys 564

Qy       1993 CAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGT 2052
              |||:::||||||:::|||||||||     :::|||:::         :::         :::
Db        565 GlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis---------TyrLys 581

Qy       2053 GGTAGTACAAGAGGAAACGTG-----------------------TTTGCATCAGTTGAT 2088
              ||||||:::||||||                              |||   :::   |||
Db        582 GlySerSerArgGlyArgSerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAsp 601

Qy       2089 ACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                 |||:::      |||      :::::::::   ||||||||||||||||||
Db        602 TyrArgGln---SerSerGlyAlaSerSerSerSerPheSerSerSerArgAla 618
```

RESULT 11
AAW81501

```
ID   AAW81501 standard; Protein; 662 AA.
XX
AC   AAW81501;
XX
DT   02-FEB-1999  (first entry)
XX
DE   Dead Box X (DBX) gene long transcript amino acid sequence.
XX
KW   Non-recombining region; human; Y chromosome; X homologue; testis; DBX;
KW   infertility; sperm; gene alteration; inhibitor; Dead Box X.
XX
OS   Homo sapiens.
XX
PN   WO9846747-A2.
XX
PD   22-OCT-1998.
XX
PF   10-APR-1998;   98WO-US07115.
XX
PR   11-APR-1997;   97US-0041877.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Lahn BT,  Page DC;
XX
DR   WPI; 1998-568729/48.
DR   N-PSDB; AAV69631.
XX
PT   Novel genes in the non-combining region of Y chromosome - useful to
PT   diagnose if male infertility or reduced sperm count has a genetic
PT   basis
XX
PS   Disclosure; Fig 3A-B; 54pp; English.
XX
CC   This represents the amino acid sequence of the Dead Box X (DBY) gene long
CC   transcript. The invention relates to genes occurring on the non-
CC   recombining region of the human Y chromosome. The sequences fall into two
CC   classes: (1) X-homologous DNA which are expressed in many organs, having
CC   functional X homologues and (2) testis-specific DNA sequences. Y
CC   chromosomal DNA from males with known conditions such as infertility and
CC   reduced sperm count can be assessed using the invention to determine
CC   whether the condition is associated with or caused by the occurrence of
CC   the gene or gene alteration. Candidate inhibitors of the enzymatic
CC   activity of the genes can be assessed using in vitro assays.
XX
SQ   Sequence   662 AA;

Alignment Scores:
Pred. No.:              7.48e-107      Length:         662
Score:                  1215.00        Matches:        293
Percent Similarity:     56.64%         Conservative:   91
Best Local Similarity:  43.22%         Mismatches:     180
```

```
Query Match:            30.73%      Indels:       114
DB:                     19          Gaps:          20

US-09-714-865-15 (1-2172) x AAW81501 (1-662)

Qy     214 GAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGT--- 270
Db      20 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 37

Qy     271 ---------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
Db      38 TyrIleProProHisLeuArgAsnArgGluAlaThrArgGly---PheTyrAspLysAsp 56

Qy     316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC---------------------GAA 354
Db      57 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76

Qy     355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
Db      77 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer------------------ 90

Qy     415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGC 468
Db      91 ---------------------------GlySerArgGlyArgPheAspAspArgGlyArg 101

Qy     469 CGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG 528
Db     102 SerAspTyrAspGlyIleGlySer------------------------------------ 109

Qy     529 CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGAT 588
Db     110 ---------------------------------------------------ArgGlyAsp 112

Qy     589 ACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAAT 645
Db     113 ------------ArgSerGlyPheGlyLysPheGluArgGlyGly------------Asn 124

Qy     646 GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
Db     125 SerArgTrpCysAspLysSerAspGluAspAspTrpSerLys------------------ 138

Qy     706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--- 762
Db     139 -------------------------------------ProLeuProProSerGluArgLeu 146

Qy     763 GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATT 822
Db     147 GluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIle 166

Qy     823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
Db     167 ProValGluAlaThrGlyAsnAsnCysProProHisIleGluSerPheSerAspValGlu 186

Qy     883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
Db     187 MetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgProThrProVal 206

Qy     943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
Db     207 GlnLysHisAlaIleProIleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThr 226

Qy    1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA 1062
Db     227 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 246

Qy    1063 ------------ATAACTGCCAGTCGT---------------TTTAAAGAGTTGCAGGAA 1095
Db     247 ProGlyGluAlaLeuArgAlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyr 266

Qy    1096 CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCC 1155
Db     267 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluGluAla 286

Qy    1156 AGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTG 1215
Db     287 ArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGlyAlaAspIle 306

Qy    1216 GGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGA 1275
Db     307 GlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArg 326

Qy    1276 CTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTG 1335
Db     327 LeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeu 346

Qy    1336 GATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT 1395
Db     347 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArgIleValGlu 366

Qy    1396 TGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCA 1455
Db     367 GlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAlaThrPhePro 386

Qy    1456 GAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT 1515
Db     387 LysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePheLeuAlaVal 405

Qy    1516 GGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTC 1575
Db     406 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpValGluGluSer 425

Qy    1576 TCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATG 1632
Db     426 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeu 445

Qy    1633 GTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAA 1692
Db     446 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGly 465

Qy    1693 ATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGA 1752
Db     466 TyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHis 485

Qy    1753 GATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTG 1812
Db     486 GlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 505

Qy    1813 GATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATAT 1872
Db     506 AspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIleGluGluTyr 525

Qy    1873 GTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTT 1932
Db     526 ValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 545

Qy    1933 GATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAA 1992
Db     546 Asn---GluArgAsnIleAsnIleThrLysAspLeuAspLeuLeuValGluAlaLys 564

Qy    1993 CAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGT 2052
Db     565 GlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis---------TyrLys 581

Qy    2053 GGTAGTACAAGAGGAAACGTG------------------------TTTGCATCAGTTGAT 2088
Db     582 GlySerSerArgGlyArgSerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAsp 601

Qy    2089 ACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
Db     602 TyrArgGln---SerSerGlyAlaSerSerSerSerPheSerSerSerArgAla 618

RESULT 12
AAY96483
ID   AAY96483 standard; Protein; 662 AA.
XX
AC   AAY96483;
```

```
XX
DT   12-SEP-2000  (first entry)
XX
DE   Human DBX1, an RNA helicase homologue.
XX
KW   Topoisomerase III-alpha; interacting protein; TIII-a IP; RNA helicase;
KW   DBX1; nucleosome; supercoiling; chromosome segregation; recombination;
KW   stabilization; cell division; apoptosis; cell cycle regulation;
KW   cytotoxic; anti-tumour; DBX1.
XX
OS   Homo sapiens.
XX
PN   WO200032768-A1.
XX
PD   08-JUN-2000.
XX
PF   29-NOV-1999;  99WO-FR02952.
XX
PR   30-NOV-1998;  98FR-0015081.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Fournier A,  Goulaouic H,  Riou J;
XX
DR   WPI; 2000-412316/35.
DR   N-PSDB; AAA29207.
XX
PT   New nucleic acid encoding ligand for topoisomerase III alpha, useful
PT   for inhibiting the enzyme and in drug screening, e.g. for potential
PT   anticancer agents
XX
PS   Disclosure; Page 53-56; 68pp; French.
XX
CC   This protein is DBX1, which shows homology with RNA helicases but the
CC   activity of a helicase has never been demonstrated and its function has
CC   not yet been identified. DBX1 possesses the 8 characteristic motifs of
CC   the helicases of the family "DEAD". In particular, it appears to be part
CC   of the sub-family represented by helicase PL10. The DBX1 gene is
CC   situated on the X chromosome and its homologue, which is situated on the
CC   Y chromosome has 91 percent identity with a novel protein coding
CC   sequence. The novel protein (see AAY96482) is a human topoisomerase
CC   III-alpha interacting protein (TIII-a IP). TIII-a IP has structural
CC   features in common with RNA helicases (e.g. DBX1), which are involved in
CC   destruction of the nucleosome structure, supercoiling of DNA, segregation
CC   of newly replicated chromosomes, and recombination and stabilization of
CC   the genome. Agents that alter interaction between TIII-a IP and TIII-a
CC   therefore modulate cell division, replication, transcription,
CC   translation, splicing and DNA recombination or repair, so may slow cell
CC   growth, block the cell cycle or induce apoptosis. Antibodies and ligands
CC   of TIII-a IP are used to prevent, treat or alleviate diseases that
CC   involve abnormal regulation of the cell cycle, i.e. they are potential
CC   anti-tumour (cytotoxic) agents.
XX
SQ   Sequence   662 AA;

Alignment Scores:
Pred. No.:               7.48e-107     Length:         662
Score:                   1215.00       Matches:        293
Percent Similarity:      56.64%        Conservative:   91
Best Local Similarity:   43.22%        Mismatches:     180
Query Match:             30.73%        Indels:         114
DB:                      21            Gaps:           20

US-09-714-865-15 (1-2172) x AAY96483 (1-662)

Qy        214 GAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTGGAGTTGGAAAGAGT--- 270
              :::  |||     ||||||  |||      ||||||          |||
Db         20 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 37

Qy        271 --------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
                            ||||||    :::             |||  |||   |||
Db         38 TyrIleProProHisLeuArgAsnArgGluAlaThrArgGly---PheTyrAspLysAsp 56

Qy        316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------------------GAA 354
              |||||||||:::                 |||
Db         57 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
              |||:::      ::::::     ||||||  ||||||
Db         77 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer----------------- 90

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGC 468
                                         |||  ||||||  |||     |||||
Db         91 --------------------------GlySerArgGlyArgPheAspAspArgGlyArg 101

Qy        469 CGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG 528
                          |||:::|||||
Db        102 SerAspTyrAspGlyIleGlySer------------------------------------ 109

Qy        529 CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGAT 588
                                                                    ||||||
Db        110 -------------------------------------------------ArgGlyAsp 112

Qy        589 ACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAAT 645
                          |||||||||  |||     |||||||||||           |||
Db        113 ------------ArgSerGlyPheGlyLysPheGluArgGlyGly-----------Asn 124

Qy        646 GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
                           |||  ::::::  |||
Db        125 SerArgTrpCysAspLysSerAspGluAspAspTrpSerLys----------------- 138

Qy        706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--- 762
                                                  |||  |||||| · |||
Db        139 ----------------------------------ProLeuProProSerGluArgLeu 146

Qy        763 GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATT 822
              |||      :::|||:::          ||||||||||||||:::|||||||||   |||
Db        147 GluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIle 166

Qy        823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                 ||||||   :::|||::::::  ||||||  |||    :::|||    :::
Db        167 ProValGluAlaThrGlyAsnAsnCysProProHisIleGluSerPheSerAspValGlu 186

Qy        883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
              :::   :::   :::         ||||||        ||||||:::   |||||||||
Db        187 MetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgProThrProVal 206

Qy        943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
              ||||||::::::|||||||||||||            ||||||||||||||||||||||||
Db        207 GlnLysHisAlaIleProIleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThr 226

Qy       1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA 1062
              ||||||||||||||||||||||||||||||||||||||||||:::   :::     ||||||
Db        227 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 246

Qy       1063 ------------ATAACTGCCAGTCGT--------------TTTAAAGAGTTGCAGGAA 1095
                          :::   |||  :::              :::      |||
Db        247 ProGlyGluAlaLeuArgAlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyr 266

Qy       1096 CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCC 1155
              |||      ::::::::||||||||||||||||||      |||||||||   ||||||
Db        267 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluGluAla 286

Qy       1156 AGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTG 1215
              ||||||||||||:::   :::   ||||||      |||:::||||||||||     :::
Db        287 ArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGlyAlaAspIle 306

Qy       1216 GGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGA 1275
              |||      ||||||   :::   :::||||||::::::|||   |||||||||||||||
Db        307 GlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArg 326

Qy       1276 CTGATGGATATCATAGGCAAAGAAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTG 1335
              |||:::|||::::::   :::   ||||||||||||        |||||||||||||||
Db        327 LeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeu 346

Qy       1336 GATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT 1395
```

```
                 ||||||||||||||||||||||||||||||     |||:::::::::::::::::
Db         347 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArgIleValGlu 366

Qy        1396 TGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCA 1455
                         |||||  |||     |||  |||:::|||||||||||||||||||
Db         367 GlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAlaThrPhePro 386

Qy        1456 GAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT 1515
               :::||||||||  |||||  :::||||||        |||:::|||:::|||||
Db         387 LysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePheLeuAlaVal 405

Qy        1516 GGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTC 1575
               |||:::||||||              ::::::  |||  |||:::  |||   :::
Db         406 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpValGluGluSer 425

Qy        1576 TCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATG 1632
                  ||||||    |||:::::::::|||          |||   |||    |||:::
Db         426 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeu 445

Qy        1633 GTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAA 1692
               ||||||||||||||||||  ||||||               ||||||    |||
Db         446 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGly 465

Qy        1693 ATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGA 1752
                  :::   ||||||||||||||||||    ||||||:::||||||:::||||||
Db         466 TyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHis 485

Qy        1753 GATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTG 1812
                  ||||||  ||||||    |||:::||||||||||||:::||||||||||||||||
Db         486 GlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 505

Qy        1813 GATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATAT 1872
               ||||||  ||||||:::||||||||||||||||||||||||       |||:::||||||
Db         506 AspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIleGluGluTyr 525

Qy        1873 GTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTT 1932
               |||||||||||||||||||||||||   ||||||   |||   |||   |||||||||
Db         526 ValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 545

Qy        1933 GATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAA 1992
               :::    |||   :::   ::::::   :::   |||:::    :::|||    :::|||:::
Db         546 Asn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuValGluAlaLys 564

Qy        1993 CAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGT 2052
               |||:::||||||:::|||||||||     :::|||:::     :::         :::
Db         565 GlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis---------TyrLys 581

Qy        2053 GGTAGTACAAGAGGAAACGTG----------------------TTTGCATCAGTTGAT 2088
               ||||||:::||||||                           |||    :::   |||
Db         582 GlySerSerArgGlyArgSerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAsp 601

Qy        2089 ACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                  |||:::      |||        :::::::::   ||||||||||||||||
Db         602 TyrArgGln---SerSerGlyAlaSerSerSerSerPheSerSerSerArgAla 618
```

RESULT 13
AAW81503

```
ID   AAW81503 standard; Protein; 660 AA.
XX
AC   AAW81503;
XX
DT   02-FEB-1999  (first entry)
XX
DE   Dead Box Y (DBY) gene product.
XX
KW   Non-recombining region; human; Y chromosome; X homologue; testis; DBY;
KW   infertility; sperm; gene alteration; inhibitor; Dead Box Y.
XX
OS   Homo sapiens.
XX
PN   WO9846747-A2.
XX
PD   22-OCT-1998.
XX
PF   10-APR-1998;   98WO-US07115.
XX
PR   11-APR-1997;   97US-0041877.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Lahn BT,  Page DC;
XX
DR   WPI; 1998-568729/48.
DR   N-PSDB; AAV69633.
XX
PT   Novel genes in the non-combining region of Y chromosome - useful to
PT   diagnose if male infertility or reduced sperm count has a genetic
PT   basis
XX
PS   Claim 6; Fig 3A-B; 54pp; English.
XX
CC   This represents the amino acid sequence of the Dead Box Y (DBY) gene
CC   product. The invention relates to genes occurring on the non-recombining
CC   region of the human Y chromosome. The sequences fall into two classes:
CC   (1) X-homologous DNA which are expressed in many organs, having
CC   functional X homologues and (2) testis-specific DNA sequences. Y
CC   chromosomal DNA from males with known conditions such as infertility and
CC   reduced sperm count can be assessed using the invention to determine
CC   whether the condition is associated with or caused by the occurrence of
CC   the gene or gene alteration. Candidate inhibitors of the enzymatic
CC   activity of the genes can be assessed using in vitro assays.
XX
SQ   Sequence   660 AA;

Alignment Scores:
Pred. No.:               9.55e-104    Length:         660
Score:                   1182.50      Matches:        278
Percent Similarity:      58.06%       Conservative:   93
Best Local Similarity:   43.51%       Mismatches:     167
Query Match:             29.91%       Indels:         101
DB:                      19           Gaps:           18

US-09-714-865-15 (1-2172) x AAW81503 (1-660)

Qy         301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------- 351
                  |||   |||    ||||||||||||:::                      |||
Db          52 PheHisAspLysAspSerSerGlyTrpSerCysSerLysAspLysAspAlaTyrSerSer 71

Qy         352 ---------GAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
                             |||:::      :::          ||||||:::||||||
Db          72 PheGlySerArgAspSerArgGlyLysProGlyTyrPheSerGluArgGlySer------ 89

Qy         403 GATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC--- 459
                                                      |||   ||||||   |||
Db          90 ---------------------------------------GlySerArgGlyArgPheAsp 96

Qy         460 ---CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGAC 516
                  ||||||                              :::|||   |||
Db          97 AspArgGly------------------------------ArgSerAspTyrAsp------ 104

Qy         517 GAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACA 576
                                                                  |||
Db         105 ---------------------------------------------------GlyIle 106

Qy         577 GGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTAC 633
               ||||||             :::   |||   |||   |||      ||||||   |||:::
Db         107 GlyAsn------------ArgGluArgProGlyPheGlyArgPheGluArgSerGlyHis 122

Qy         634 AAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCA 693
                     ::::::             |||      ::::::   |||
Db         123 SerArgTrpCysAspLys------------SerValGluAspAspTrpSerLys------ 136

Qy         694 GAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCA 753
                                                              |||   ||||||
Db         137 ------------------------------------------------ProLeuProPro 140
```

```
Qy        754 CCTGAGGAT---GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAA 810
Db        141 SerGluArgLeuGluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLys 160

Qy        811 TACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTT 870
Db        161 TyrAspAspIleProValGluAlaThrGlySerAsnCysProProHisIleGluAsnPhe 180

Qy        871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
Db        181 SerAspIleAspMetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArg 200

Qy        931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
Db        201 ProThrProValGlnLysHisAlaIleProIleIleLysGlyLysArgAspLeuValAla 220

Qy        991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG 1050
Db        221 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIle 240

Qy       1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA--------------- 1095
Db        241 TyrThrAspGly---ProGlyGluAlaLeuLysAlaValLysGluAsnGlyArgTyrGly 259

Qy       1096 ---------------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
Db        260 ArgArgLysGlnTyrProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGln 279

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
Db        280 IleTyrGluGluAlaArgLysPheSerTyrArgSerArgValArgProCysValValTyr 299

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
Db        300 GlyGlyAlaAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuVal 319

Qy       1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
Db        320 AlaThrProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCys 339

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
Db        340 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIle 359

Qy       1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
Db        360 ArgArgIleValGluGlnAspThrMetProProLysGlyValArgHisThrMetMetPhe 379

Qy       1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
Db        380 SerAlaThrPheProLysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyr 398

Qy       1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
Db        399 IlePheLeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValVal 418

Qy       1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG--- 1617
Db        419 TrpValGluAspLeuAspLysArgSerPheLeuLeuAspIleLeuGlyAlaThrGlySer 438

Qy       1618 GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTT 1677
Db        439 AspSerLeuThrLeuValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPhe 458

Qy       1678 CTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGG 1737
Db        459 LeuTyrHisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArg 478

Qy       1738 GAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTA 1797
Db        479 GluGluAlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaVal 498

Qy       1798 GCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCT 1857
Db        499 AlaAlaArgGlyLeuAspIleSerAsnValArgHisValIleAsnPheAspLeuProSer 518

Qy       1858 ACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGA 1917
Db        519 AspIleGluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeu 538

Qy       1918 GCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTA 1977
Db        539 AlaThrSerPhePheAsn---GluLysAsnMetAsnIleThrLysAspLeuLeuAspLeu 557

Qy       1978 TTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATAC 2037
Db        558 LeuValGluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis 577

Qy       2038 ATTCCTGGCTTCAGT---GGTAGTACAAGAGGAAACGTGTTTGCA------------TCA 2082
Db        578 TyrLysGlyGlySerArgGlyArgSerLysSerAsnArgPheSerGlyGlyPheGlyAla 597

Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
Db        598 ArgAspTyrArgGlnSerSerGlySer---SerSerSerGlyPheGlyAlaSerArg 615
```

```
RESULT 14
AAG17433
ID   AAG17433 standard; Protein; 612 AA.
XX
AC   AAG17433;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 18449.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
```

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

```
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

Alignment Scores:
Pred. No.:               8.79e-96     Length:         612
Score:                   1099.00      Matches:        252
Percent Similarity:      57.97%       Conservative:   97
Best Local Similarity:   41.86%       Mismatches:     167
Query Match:             27.79%       Indels:         86
DB:                      21           Gaps:           15

US-09-714-865-15 (1-2172) x AAG17433 (1-612)

Qy        313 GATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAAC 372
              |||   :::|||:::      :::    :::                |||:::
Db         43 AspHisAlaGlyTyrGlyGlyGlnProAlaGlySerArgTrpAlaProProSerSerGly 62

Qy        373 AGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCA 432
                 |||   :::      ||||||||||||:::
Db         63 GlyGlyGlyAlaSerGlyGlyGlyTyrArgAsnAsp------------------------ 74

Qy        433 TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGT 492
                       |||||||||       :::               |||:::|||   |||
Db         75 ---------GlyGlyArgThrGlyTyr---------------GlyTyrGlyAlaGly--- 85

Qy        493 CCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCT 552
                                                          ||||||
Db         86 ------------------------------------------GlyGly------------ 87

Qy        553 AGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGT 612
                                |||   |||   |||                    |||
Db         88 ------------------GlyGlyGlyGlyGly--------------------GlyTrp 94

Qy        613 GGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAG 672
                :::         ||||||:::              ||||||
Db         95 AsnAsnArgSerGlyGlyTrpAspArgArgGluArgGluVal------------------ 108

Qy        673 AATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTG 732
              |||   :::      :::||||||
Db        109 AsnProPheGlyAspAspAlaGlu------------------------------------ 116

Qy        733 ACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACA 792
                    :::   |||         |||   |||:::                    |||
Db        117 ------LeuGluProValPheThrGluGlnGluAsn--------------------Thr 127

Qy        793 GGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCA 852
              ||||||||||||||   |||:::   |||   ||||||   ||||||   |||   |||
Db        128 GlyIleAsnPheAspAlaTyrGluAspIleProValGluThrSerGlyGlyAspValPro 147

Qy        853 CCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCT 912
              |||   :::   ||||||   :::   :::|||         ||||||   ||||||
Db        148 ProProValAsnThrPheAlaAspIleAspLeuGlyAspAlaLeuAsnLeuAsnIleArg 167

Qy        913 AAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCA 972
              :::      |||   :::   ||||||||||||:::::::::|||||||||:::||||||
Db        168 ArgCysLysTyrValArgProThrProValGlnArgHisAlaIleProIleLeuLeuAla 187

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 GluArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheCysPhe 207

Qy       1033 CCAATTTTGGCTCATATGATGCATGAT---GGAATAACTGCCAGTCGTTTTAAAGAGTTG 1089
              ||||||::::::   :::|||   |||        :::            |||
Db        208 ProIleIleSerGlyIleMetLysAspGlnHisValGluArgProArgGlySerArgAla 227

Qy       1090 CAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149
                 |||      :::|||::::::||||||||||||||||      ||||||:::
Db        228 ValTyrProPheAlaValIleLeuSerProThrArgGluLeuAlaCysGlnIleHisAsp 247

Qy       1150 GAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACC 1209
              ||||||:::|||||||||:::   |||   |||:::   ||||||   ||||||||||||
Db        248 GluAlaLysLysPheSerTyrGlnThrGlyValLysValValValAlaTyrGlyGlyThr 267

Qy       1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269
                 :::         :::|||::::::   :::||||||:::||||||   |||||||||
Db        268 ProIleHisGlnGlnLeuArgGluLeuGluArgGlyCysAspIleLeuValAlaThrPro 287

Qy       1270 GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA 1329
              |||||||||   |||::::::   :::   ::::::   ::::::   |||::::::|||
Db        288 GlyArgLeuAsnAspLeuLeuGluArgAlaArgValSerMetGlnMetIleArgPheLeu 307

Qy       1330 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTA 1389
                 ||||||||||||||||||||||||||||||||||||   |||:::::::::|||:::
Db        308 AlaLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgLysIle 327

Qy       1390 ATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACT 1449
              :::               ||||||   :::      |||||||||::::::||||||||||||
Db        328 ValGluGlnMetAspMetProProArgGlyValArgGlnThrMetLeuPheSerAlaThr 347

Qy       1450 TTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTT 1509
              ||||||   :::||||||||||||||||||::::||:::   ||||||||:::|||:::
Db        348 PheProSerGlnIleGlnArgLeuAlaAlaAspPheMet---SerAsnTyrIlePheLeu 366

Qy       1510 GCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGC 1569
              ||||||||:::||||||   :::            :::   |||   |||      |||
Db        367 AlaValGlyArgValGlySerSerThrAspLeuIleThrGlnArgValGluPheValGln 386

Qy       1570 CAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTG---------CGAAACATAGGGGAT 1620
              :::      ||||||   |||:::::::::|||            |||         |||
Db        387 GluSerAspLysArgSerHisLeuMetAspLeuLeuHisAlaGlnArgGluThrGlnAsp 406

Qy       1621 GAAAGA------ACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
              ::::::         |||:::||||||||||||||||||:::   ||||||
Db        407 LysGlnSerLeuThrLeuValPheValGluThrLysArgGlyAlaAspThrLeuGluAsn 426

Qy       1675 TTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
              :::||||||      :::         |||||||||||||||||||||   |||:::|||
Db        427 TrpLeuCysMetAsnGluPheProAlaThrSerIleHisGlyAspArgThrGlnGlnGlu 446

Qy       1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
              ||||||   ||||||      |||:::   |||:::   |||:::|||||||||||||
Db        447 ArgGluValAlaLeuArgSerPheLysThrGlyArgThrProIleLeuValAlaThrAsp 466

Qy       1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
              ||||||||||||||||||||||||      :::|||   ||||||:::||||||||||||||
Db        467 ValAlaAlaArgGlyLeuAspIleProHisValAlaHisValValAsnPheAspLeuPro 486

Qy       1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
              :::   |||||:::||||||||||||||||||||||||||||   |||   :::|||
Db        487 AsnAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAlaGlyLysSerGly 506

Qy       1915 AGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAA 1974
                 |||   :::||||||:::   |||::::::         ||||||:::   |||   :::
Db        507 IleAlaThrAlaPhePheAsn---GluAsnAsnAlaGlnLeuAlaArgSerLeuAlaGlu 525

Qy       1975 GTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACA 2034
              ::::::      :::|||   |||:::||||||   ||||||         |||
Db        526 LeuMetGlnGluAlaAsnGlnGluValProGluTrpLeuThrArgTyrAlaSerArgAla 545

Qy       2035 TACATTCCTGGCTTCAGTGGTAGTACAAGA---GGAAACGTGTTTGCATCAGTTGATACC 2091
                    |||      |||      |||   |||         |||         |||
Db        546 SerPheGlyGly------GlyLysLysArgSerGlyGlyArgPheGlyGlyArgAspPhe 563

Qy       2092 AGAAAG 2097
              |||:::
Db        564 ArgArg 565

RESULT 15
AAG20655
ID   AAG20655 standard; Protein; 612 AA.
XX
AC   AAG20655;
```

```
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 22931.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
```

```
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

Alignment Scores:
Pred. No.:                8.79e-96      Length:         612
Score:                    1099.00       Matches:        252
Percent Similarity:       57.97%        Conservative:   97
Best Local Similarity:    41.86%        Mismatches:     167
Query Match:              27.79%        Indels:         86
DB:                       21            Gaps:           15

US-09-714-865-15 (1-2172) x AAG20655 (1-612)

Qy        313 GATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAAC 372
              |||   :::|||:::      :::    :::                 |||:::
Db         43 AspHisAlaGlyTyrGlyGlyGlnProAlaGlySerArgTrpAlaProProSerSerGly 62

Qy        373 AGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCA 432
                 |||   :::      ||||||||||||:::
Db         63 GlyGlyGlyAlaSerGlyGlyGlyTyrArgAsnAsp------------------------ 74

Qy        433 TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGT 492
                       |||||||||      :::                |||:::|||   |||
Db         75 ---------GlyGlyArgThrGlyTyr---------------GlyTyrGlyAlaGly--- 85

Qy        493 CCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCT 552
                                                        ||||||
Db         86 ----------------------------------------GlyGly------------ 87

Qy        553 AGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGT 612
                               |||   |||   |||                    |||
Db         88 -----------------GlyGlyGlyGlyGly--------------------GlyTrp 94

Qy        613 GGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAG 672
                 :::      ||||||:::                 ||||||
Db         95 AsnAsnArgSerGlyGlyTrpAspArgArgGluArgGluVal------------------ 108

Qy        673 AATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTG 732
              |||   :::      :::||||||
Db        109 AsnProPheGlyAspAspAlaGlu------------------------------------ 116

Qy        733 ACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACA 792
                    :::   |||         |||   |||:::                   |||
Db        117 ------LeuGluProValPheThrGluGlnGluAsn--------------------Thr 127

Qy        793 GGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCA 852
              |||||||||||||||   |||:::   |||   ||||||   ||||||   |||   |||
Db        128 GlyIleAsnPheAspAlaTyrGluAspIleProValGluThrSerGlyGlyAspValPro 147

Qy        853 CCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCT 912
              |||   :::   ||||||   :::   :::|||         ||||||   ||||||
Db        148 ProProValAsnThrPheAlaAspIleAspLeuGlyAspAlaLeuAsnLeuAsnIleArg 167

Qy        913 AAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCA 972
              :::      |||   :::   ||||||||||||:::::::::|||||||||:::||||||
Db        168 ArgCysLysTyrValArgProThrProValGlnArgHisAlaIleProIleLeuLeuAla 187

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 GluArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheCysPhe 207

Qy       1033 CCAATTTTGGCTCATATGATGCATGAT---GGAATAACTGCCAGTCGTTTTAAAGAGTTG 1089
              ||||||::::::   :::|||   |||         :::         |||
Db        208 ProIleIleSerGlyIleMetLysAspGlnHisValGluArgProArgGlySerArgAla 227

Qy       1090 CAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149
                    |||      :::|||:::::::|||||||||||||||      ||||||:::
Db        228 ValTyrProPheAlaValIleLeuSerProThrArgGluLeuAlaCysGlnIleHisAsp 247

Qy       1150 GAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACC 1209
              ||||||:::||||||||||:::   |||   |||:::   ||||||   ||||||||||||
Db        248 GluAlaLysLysPheSerTyrGlnThrGlyValLysValValValAlaTyrGlyGlyThr 267

Qy       1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269
                 :::         :::|||::::::   :::||||||:::||||||   |||||||||
Db        268 ProIleHisGlnGlnLeuArgGluLeuGluArgGlyCysAspIleLeuValAlaThrPro 287

Qy       1270 GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA 1329
              |||||||||   |||::::::   :::   ::::::   ::::::   |||::::::|||
Db        288 GlyArgLeuAsnAspLeuLeuGluArgAlaArgValSerMetGlnMetIleArgPheLeu 307

Qy       1330 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTA 1389
                 ||||||||||||||||||||||||||||||||||||   |||:::::::::|||:::
Db        308 AlaLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgLysIle 327

Qy       1390 ATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACT 1449
              :::                  ||||||   :::      |||||||||::::::||||||||||||
Db        328 ValGluGlnMetAspMetProProArgGlyValArgGlnThrMetLeuPheSerAlaThr 347

Qy       1450 TTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTT 1509
              ||||||   :::|||||||||||||||||:::|||:::   |||||||||:::|||:::
Db        348 PheProSerGlnIleGlnArgLeuAlaAlaAspPheMet---SerAsnTyrIlePheLeu 366

Qy       1510 GCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGC 1569
              |||||||||:::||||||      :::         :::   |||   |||      |||
Db        367 AlaValGlyArgValGlySerSerThrAspLeuIleThrGlnArgValGluPheValGln 386
```

```
Qy       1570 CAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTG---------CGAAACATAGGGGAT 1620
              :::      ||||||     |||:::::::::|||          |||         |||
Db        387 GluSerAspLysArgSerHisLeuMetAspLeuLeuHisAlaGlnArgGluThrGlnAsp 406

Qy       1621 GAAAGA------ACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
              ::::::      |||:::||||||||||||||||||:::    ||||||
Db        407 LysGlnSerLeuThrLeuValPheValGluThrLysArgGlyAlaAspThrLeuGluAsn 426

Qy       1675 TTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
              :::||||||       :::         ||||||||||||||||||||   |||:::|||
Db        427 TrpLeuCysMetAsnGluPheProAlaThrSerIleHisGlyAspArgThrGlnGlnGlu 446

Qy       1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
              ||||||   ||||||      |||:::   |||:::   |||:::||||||||||||
Db        447 ArgGluValAlaLeuArgSerPheLysThrGlyArgThrProIleLeuValAlaThrAsp 466

Qy       1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
              ||||||||||||||||||||||||    :::|||   ||||||:::||||||||||||||
Db        467 ValAlaAlaArgGlyLeuAspIleProHisValAlaHisValValAsnPheAspLeuPro 486

Qy       1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
              :::    ||||||:::||||||||||||||||||||||||||||||   |||   :::|||
Db        487 AsnAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAlaGlyLysSerGly 506

Qy       1915 AGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAA 1974
                 |||   :::||||||:::   |||::::::         ||||||:::   |||   :::
Db        507 IleAlaThrAlaPhePheAsn---GluAsnAsnAlaGlnLeuAlaArgSerLeuAlaGlu 525

Qy       1975 GTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACA 2034
              ::::::   :::|||   |||:::||||||   ||||||         |||
Db        526 LeuMetGlnGluAlaAsnGlnGluValProGluTrpLeuThrArgTyrAlaSerArgAla 545

Qy       2035 TACATTCCTGGCTTCAGTGGTAGTACAAGA---GGAAACGTGTTTGCATCAGTTGATACC 2091
                       |||      |||      |||   |||      |||         |||
Db        546 SerPheGlyGly------GlyLysLysArgSerGlyGlyArgPheGlyGlyArgAspPhe 563

Qy       2092 AGAAAG 2097
              |||:::
Db        564 ArgArg 565
```

Search completed: June 10, 2003, 16:38:28
Job time : 124.5 secs

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          June 10, 2003, 06:09:16 ; Search time 2771 Seconds
                                           (without alignments)
                                           12694.550 Million cell updates/sec

Title:           US-09-714-865-15
Perfect score:   2172
Sequence:        1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                        Description
--------------------------------------------------------------------------------
     1  2150.4    99.0   2523 11  BC030638                  BC030638 Homo sapi
     2  1528.4    70.4   2759 11  AK014844                  AK014844 Mus muscu
     3     683    31.4    798 12  BG717110                  BG717110 602689191
     4   640.8    29.5    822 13  BI464267                  BI464267 603203379
     5   631.8    29.1    792 14  BQ429126                  BQ429126 AGENCOURT
     6   624.8    28.8    747 12  BG719449                  BG719449 602690172
     7   620.4    28.6    682 12  BG718234                  BG718234 602696209
     8   620.4    28.6    693 13  BI560849                  BI560849 603254018
     9   615.6    28.3    830 13  BI462207                  BI462207 603205520
    10   612.4    28.2    918 13  BI561001                  BI561001 603254391
    11   610.4    28.1    846 12  BG718634                  BG718634 602696739
    12   597.4    27.5    699 12  BG771822                  BG771822 602720504
    13   596.6    27.5    758 13  BI559479                  BI559479 603252992
    14   594.8    27.4    861 13  BI461393                  BI461393 603206861
    15   593.4    27.3    657 12  BG717812                  BG717812 602693829
    16   584.6    26.9    743 13  BI463502                  BI463502 603203596
    17   581.8    26.8    884 13  BI560499                  BI560499 603254476
    18   564.4    26.0    651 12  BG717354                  BG717354 602689688
    19   559.8    25.8    700 12  BG718861                  BG718861 602699337
    20   544.4    25.1    847 13  BI562624                  BI562624 603256559
    21   542.6    25.0    810 13  BI463846                  BI463846 603202572
    22   524.8    24.2    841 13  BI459115                  BI459115 603199485
    23     504    23.2    917 13  BI463447                  BI463447 603204483
    24   483.6    22.3    546 13  BM539751                  BM539751 hb13g02.g
    25   449.8    20.7    670 14  BQ443738                  BQ443738 UI-M-EW0-
    26   433.8    20.0    482  9  AL042306                  AL042306 DKFZp434M
    27   420.2    19.3    780 10  AW153705                  AW153705 fi24g05.y
    28     389    17.9    726 12  BF611759                  BF611759 de88d02.y
    29     371    17.1   4626 11  BC007668                  BC007668 Homo sapi
    30   368.8    17.0    604 13  BI460628                  BI460628 603201271
    31     339    15.6    632  9  AL785068                  AL785068 AL785068
    32     333    15.3    335  9  AA383535                  AA383535 EST96928
c   33   310.4    14.3    555  9  AA399611                  AA399611 zt93a07.s
    34   298.4    13.7    491  9  AA398976                  AA398976 zt93a07.r
    35   296.2    13.6    537 14  BQ520442                  BQ520442 NISC_nl04
c   36   290.8    13.4    430  9  AI843771                  AI843771 UI-M-AN1-
c   37   288.4    13.3    555  9  AI217144                  AI217144 qf47d11.x
    38     270    12.4    533 10  AW634656                  AW634656 bl23d02.w
    39   258.4    11.9   1084 13  BM462228                  BM462228 AGENCOURT
    40   247.6    11.4    976 17  CNS071VS                  AL425438 clone BA0
    41   237.4    10.9    457 13  BJ073747                  BJ073747 BJ073747
    42   236.6    10.9    819  9  AA791735                  AA791735 vu08b08.r
c   43   235.4    10.8    347  9  AI953070                  AI953070 wq49h07.x
    44   222.6    10.2    637 13  BI461238                  BI461238 603206562
    45   220.2    10.1    816  9  AU132682                  AU132682 AU132682

                              ALIGNMENTS

RESULT 1
BC030638
LOCUS       BC030638                2523 bp    mRNA    linear   HTC 21-MAY-2002
DEFINITION  Homo sapiens, VASA protein, clone IMAGE:4822528, mRNA.
ACCESSION   BC030638
VERSION     BC030638.1  GI:21040391
KEYWORDS    HTC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2523)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
```

```
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 33 Row: d Column: 8
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 9507236
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..2523
                     /organism="Homo sapiens"
                     /db_xref="LocusID:54514"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4822528"
                     /tissue_type="Testis"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
BASE COUNT      780 a    437 c    616 g    690 t
ORIGIN

  Query Match             99.0%;  Score 2150.4;  DB 11;  Length 2523;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2165;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 157

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 217

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 277

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 337

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 397

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 457

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 517

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 577

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 637

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 697

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 757

Qy        661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 817

Qy        721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 877

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 937

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 997

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 1057

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db       1058 ATCATACTTGCAGGACGAGATTTGATGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGCG 1117

Qy       1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1118 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1177

Qy       1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1178 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1237

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1238 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1297

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1298 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1357

Qy       1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1358 GCTACTCC-GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1416

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1476

Qy       1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1536

Qy       1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1596

Qy       1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1656

Qy       1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1657 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1716

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db       1717 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTATTGCAACTTTTCTT 1776

Qy       1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db       1777 TGTCAAGAAAAAATATCAACTACAAGTATTCATGGTGATCGGGAACAGAGAGAGCGGGAG 1836

Qy       1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1837 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1896

Qy       1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
```

```
                 |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db          1897 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAGTTTTGATCTTCCTTCTACC 1956

Qy          1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1957 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 2016

Qy          1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2017 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 2076

Qy          1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2077 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2136

Qy          2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2137 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2196

Qy          2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
                 ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db          2197 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACAAGCTCCCAATCCAGTAGATGAT 2256

Qy          2161 GAGTCATGGGAT 2172
                 ||||||||||||
Db          2257 GAGTCATGGGAT 2268


RESULT 2
AK014844
LOCUS       AK014844                2759 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4921508G22:DEAD
            (aspartate-glutamate-alanine-aspartate) box polypeptide 4, full
            insert sequence.
ACCESSION   AK014844
VERSION     AK014844.1  GI:12852921
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:4921508G22.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
```

QP-01872

```
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2759)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised longer than 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..2759
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4921508G22"
                     /db_xref="MGD:MGI:1910815"
                     /db_xref="taxon:10090"
                     /clone="4921508G22"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
```

```
     gene            1..2759
                     /gene="Ddx4"
     CDS             42..2150
                     /gene="Ddx4"
                     /note="DEAD (aspartate-glutamate-alanine-aspartate) box
                     polypeptide 4
                     data source:MGD, source key:MGI:102670, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB29578.1"
                     /db_xref="GI:12852922"
                     /db_xref="MGD:MGI:102670"
                     /translation="MGDEDWEAEILKPHVSSYVPVFEKDKYSSGANGDTFNRTSASSD
                     IGESSKKENTSTTGGFGRGKGFGNRGFLNNKFEEGDSSGFWKESNNDCEDNQTRSRGF
                     SKRGGCQDGNDSEASGPFRRGGRGSFRGCRGGFGLGRPNSESDQDQGTQRGGGLFGSR
                     KPAASDSGNGDTYQSRSGSGRGGYKGLNEEVVTGSGKNSWKSETEGGESSDSQGPKVT
                     YIPPPPPEDEDSIFAHYQTGINFDKYDTILVEVSGHDAPPAILTFEEANLCQTLNNNI
                     AKAGYTKLTPVQKYSIPIVLAGRDLMACAQTGSGKTAAFLLPILAHMMRDGITASRFK
                     ELQEPECIIVAPTRELINQIYLEARKFSFGTCVRAVVIYGGTQFGHSVRQIVQGCNIL
                     CATPGRLMDIIGKEKIGLKQVKYLVLDEADRMLDMGFGPEMKKLISCPGMPSKEQRQT
                     LLFSATFPEEIQRLAGDFLKSSYLFVAVGQVGGACRDVQQTILQVGQYSKREKLVEIL
                     RNIGDERTMVFVETKKKADFIATFLCQEKISTTSIHGDREQREREQALGDFRCGKCPV
                     LVATSVAARGLDIENVQHVINFDLPSTIDEYVHRIGRTGRCGNTGRAISFFDTDSDNH
                     LAQPLVKVLSDAQQDVPAWLEEIAFSTYVPPSFSSSTRGGAVFASVDTRKNYQGKHTL
                     NTAGISSSQAPNPVDDESWD"
     polyA_signal    2734..2739
                     /gene="Ddx4"
                     /note="putative"
     polyA_site      2759
                     /gene="Ddx4"
                     /note="putative"
BASE COUNT      836 a    483 c    653 g    787 t
ORIGIN

  Query Match             70.4%;  Score 1528.4;  DB 11;  Length 2759;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches 1743;  Conservative    0;  Mismatches  256;  Indels   21;  Gaps    4;

Qy        168 AAGTGGATTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCG 227
Db        134 AAATGGAGACACTTTTAACAGGACTTCAGCTTCATCAGATATTGGCGAGTCTAGTAAAAA 193

Qy        228 AGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTT 287
Db        194 AGAGAACACATCTACAACTGGTGGCTTTGGAAGAGGAAAGGGCTTTGGAAACAGAGGTTT 253

Qy        288 TTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGA 347
Db        254 TTTAAATAACAAGTTTGAAGAAGGTGATAGCTCTGGTTTCTGGAAAGAGTCTAATAATGA 313

Qy        348 CTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGG 407
Db        314 CTGTGAAGATAATCAGACTCGAAGCAGAGGGTTTTCCAAGCGAGGTGGCTGCCAAGATGG 373

Qy        408 AAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTG 467
Db        374 AAACGATTCAGAAGCATCAGGCCCGTTCAGAAGAGGGGGAAGAGGCAGTTTCCGAGGCTG 433

Qy        468 CCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCA 527
Db        434 CCGTGGAGGATTTGGTCTAGGAAGACCAAATAGTGAATCTGACCAAGATCAGGGGACACA 493

Qy        528 GCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGA 587
Db        494 GCGTGGTGGTGGCCTTTTTGGTTCTAGGAAACCAGCAGCAAGTGATTCAGGCAATGGTGA 553

Qy        588 TACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGA 647
Db        554 CACTTACCAAAGCAGA------AGTGGAAGTGGTCGAGGTGGTTACAAAGGTTTAAATGA 607

Qy        648 AGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAG 707
Db        608 AGAAGTAGTAACAGGTTCTGGAAAGAATTCTTGGAAGTCAGAAACTGAAGGAGGTGAAAG 667

Qy        708 TAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGA 767
Db        668 CAGTGATAGTCAAGGTCCAAAAGTGACATATATACCCCCTCCTCCACCAGAGGATGAGGA 727

Qy        768 CTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGT 827
Db        728 CTCCATCTTTGCACATTATCAGACAGGCATAAACTTTGATAAATATGACACCATACTTGT 787

Qy        828 GGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTG 887
Db        788 TGAAGTATCTGGACATGATGCACCACCGGCAATTTTGACTTTTGAAGAAGCTAATCTCTG 847

Qy        888 TCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAA 947
Db        848 TCAGACACTGAATAACAACATTGCTAAAGCTGGCTATACTAAGCTTACTCCTGTGCAGAA 907

Qy        948 ATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTC 1007
Db        908 GTACAGCATTCCCATTGTATTAGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTC 967

Qy       1008 TGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAAC 1067
Db        968 TGGGAAGACTGCAGCTTTTCTCTTGCCTATTTTGGCTCATATGATGCGGGATGGAATAAC 1027

Qy       1068 TGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGA 1127
Db       1028 TGCCAGTCGCTTTAAAGAACTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGA 1087

Qy       1128 ATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGC 1187
Db       1088 ACTGATCAACCAAATTTACTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGC 1147

Qy       1188 TGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTG 1247
Db       1148 TGTGGTCATATATGGAGGAACCCAGTTTGGTCATTCAGTTCGACAGATAGTACAAGGGTG 1207

Qy       1248 TAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGG 1307
Db       1208 TAACATATTGTGTGCTACTCCAGGGAGGCTGATGGACATCATAGGCAAAGAAAAGATTGG 1267

Qy       1308 TCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTT 1367
Db       1268 CCTCAAACAAGTCAAGTACTTAGTTTTGGATGAAGCTGATCGAATGTTGGATATGGGTTT 1327

Qy       1368 TGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCA 1427
Db       1328 TGGACCAGAGATGAAGAAACTAATTTCTTGTCCAGGAATGCCATCAAAGGAACAACGCCA 1387

Qy       1428 AACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTT 1487
Db       1388 AACCCTTTTATTCAGTGCTACTTTTCCAGAAGAAATCCAGAGGTTGGCTGGGGACTTTCT 1447

Qy       1488 AAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCA 1547
Db       1448 AAAGTCCAGTTACTTGTTTGTCGCTGTTGGGCAAGTGGGAGGAGCTTGCAGAGATGTTCA 1507

Qy       1548 GCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCG 1607
Db       1508 GCAGACGATCCTTCAAGTTGGCCAGTATTCAAAAAGAGAAAAGCTTGTTGAGATTCTACG 1567

Qy       1608 AAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTAC 1667
Db       1568 AAACATAGGTGATGAAAGAACTATGGTCTTTGTTGAAACCAAGAAAAAAGCCGATTTCAT 1627

Qy       1668 TGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACA 1727
Db       1628 TGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATTCATGGTGATCGGGAGCA 1687

Qy       1728 GAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGC 1787
Db       1688 GAGGGAGAGAGAGCAAGCTCTTGGAGATTTTCGCTGTGGAAAGTGCCCAGTTCTTGTTGC 1747
```

```
Qy     1788 TACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGA 1847
            ||||||||| |||||||||||||| |||||||||||||| ||||||||||||||||||||
Db     1748 TACTTCAGTGGCTGCCAGAGGGCTTGATATTGAAAATGTTCAACATGTTATCAATTTTGA 1807

Qy     1848 TCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAA 1907
             |||||||||||||||||||| ||||||||||||||||| || ||||| || ||||| ||
Db     1808 CCTTCCTTCTACCATTGATGAGTATGTTCATCGAATTGGACGCACTGGCCGCTGTGGAAA 1867

Qy     1908 TACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCT 1967
            |||||||||||| ||||| |||||||||  ||| || ||||| |||||||||||||||||
Db     1868 TACTGGCAGAGCGATTTCTTTTTTTGATACTGACTCTGATAATCATTTAGCACAGCCTCT 1927

Qy     1968 AGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTT 2027
            ||| |||||| || |||| |||||||||||||| || |||||| | ||||| ||||||||
Db     1928 AGTTAAAGTACTGTCAGACGCTCAACAGGATGTCCCCGCATGGCTAGAAGAGATTGCCTT 1987

Qy     2028 TAGTACATACAT---TCCTGGCTTCAGTGGTAGTACAAGA---GGAAACGTGTTTGCATC 2081
             ||||| ||  |   |||  |||||||| | || ||||||   ||    |||||||||||
Db     1988 CAGTACCTATGTGCCTCCCAGCTTCAGTAGCAGCACAAGAGGGGGGGCTGTGTTTGCATC 2047

Qy     2082 AGTTGATACCAGAA---------AGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTC 2132
             ||||| || || |         ||||||||  ||| ||||| ||||| ||| |||||||
Db     2048 TGTTGACACGAGGAAGAATTACCAGGGCAAGCACACGTTGAATACAGCGGGGATTTCTTC 2107

Qy     2133 TTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
            ||||| ||||||||||||||| ||||| ||||||||||||
Db     2108 TTCACAAGCTCCCAATCCAGTTGATGACGAGTCATGGGAT 2147
```

```
RESULT 3
BG717110
LOCUS       BG717110                 798 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602689191F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821503 5',
            mRNA sequence.
ACCESSION   BG717110
VERSION     BG717110.1  GI:13996297
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 798)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10728  row: e  column: 24
            High quality sequence stop: 770.
FEATURES             Location/Qualifiers
     source          1..798
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4821503"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      254 a    124 c    233 g    186 t      1 others
ORIGIN

  Query Match             31.4%;  Score 683;  DB 12;  Length 798;
  Best Local Similarity   96.4%;  Pred. No. 1.5e-181;
  Matches  720;  Conservative    0;  Mismatches   25;  Indels    2;  Gaps    2;
```

```
Qy        1 ATGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
            ||||| |||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db       52 ATGGGAGATGAAGA-TGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 110

Qy       61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      111 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 170

Qy      121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      171 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 230

Qy      181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      231 TCTGGGCGGAATTTTGAAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 290

Qy      241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      291 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 350

Qy      301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      351 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 410

Qy      361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      411 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 470

Qy      421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      471 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 530

Qy      481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      531 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 590

Qy      541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      591 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 650

Qy      601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      651 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 710

Qy      661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGT-AGTGATACTCA 719
            |||||||||||||||||| | |      |     ||||||||| |||| |||||||||||
Db      711 GGCTCTGGAAAGAATTCTGGAAGTCAGAAGCAGCAAGGAGGAGCAAGTAAGTGATACTCA 770

Qy      720 AGGACCAAAAGTGACCTACATACCCCC 746
            |||  | |||||   ||| ||  ||||
Db      771 AGGGACCAAAGTTGACTAAATTACCCC 797
```

```
RESULT 4
BI464267
LOCUS       BI464267                 822 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603203379F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269434 5',
            mRNA sequence.
ACCESSION   BI464267
VERSION     BI464267.1  GI:15254923
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 822)
```

```
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11679  row: m  column: 19
            High quality sequence stop: 714.
FEATURES             Location/Qualifiers
     source          1..822
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5269434"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      255 a    140 c    239 g    188 t
ORIGIN

  Query Match            29.5%;  Score 640.8;  DB 13;  Length 822;
  Best Local Similarity  95.4%;  Pred. No. 1.3e-169;
  Matches  725;  Conservative    0;  Mismatches   27;  Indels    8;  Gaps    6;

Qy          1 ATGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 123

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 183

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 243

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 303

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 363

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 423

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 483

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGA-AGAGGTAGTTTCCGAGGTTGCCGTGG-AGGAT 478
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||
Db        484 GCTTCAGGGCCATACAGAAGAGGTGGACAGAGGTAGTTTCCGAGGTTGCCGTGGCAGGAT 543

Qy        479 TTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTG 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 TTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTG 603

Qy        539 GCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAA 598
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCC-TTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAA 662

Qy        599 GCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAA 658
              |||||||||||||||||||||| ||| |  | |   || | ||  ||||||||||||||| 
Db        663 GCAGAAGTGGCAGTGGAAGTGACCGAAGGTGGTTACAAAGGTTTAATGAAGAAGTAATAC 722

Qy        659 CAGG---CTCTGGAAAGAATTCTTGGAAGTCAG-AAGCAGAAGGAGGAGAAAGTAGTGAT 714
              ||||   ||  | |||||||||| ||||||||| |||||||   |  ||||||||||||||
Db        723 CAGGCTTCTGGGCAAAGAATTCTGGGAAGTCAGAAAGCAGACCGCAGAGAAAGTAGTGAT 782

Qy        715 ACTCAAGG-ACCAAAAGTGACCTACATACCCCCTCCTCCA 753
              || ||||| |||||||||||||||     |||||||||||
Db        783 ACCCAAGGAACCAAAAGTGACCTAACATACCCCTCCTCCA 822


RESULT 5
BQ429126
LOCUS       BQ429126                 792 bp    mRNA    linear   EST 24-MAY-2002
DEFINITION  AGENCOURT_7901787 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6104663
            5', mRNA sequence.
ACCESSION   BQ429126
VERSION     BQ429126.1  GI:21168202
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2342  row: n  column: 24
            High quality sequence stop: 496.
FEATURES             Location/Qualifiers
     source          1..792
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6104663"
                     /clone_lib="NIH_MGC_82"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
BASE COUNT      241 a    128 c    242 g    180 t      1 others
ORIGIN

  Query Match            29.1%;  Score 631.8;  DB 14;  Length 792;
  Best Local Similarity  96.9%;  Pred. No. 4.6e-167;
  Matches  665;  Conservative    0;  Mismatches   18;  Indels    3;  Gaps    2;

Qy          1 ATGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 120
```

```
Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 180

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 240

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 300

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 360

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 420

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 480

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 540

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGG- 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        541 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 600

Qy        540 CCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAG 599
              ||||||||||||||||||||| |||||||||||||||||||| |||||||||||||||||
Db        601 CCTTTTTGGTTCTAGAAGACCCGTATTAAGTGGCACAGGTAAGGGTGATACTTCTCAAAG 660

Qy        600 CAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAAC 659
              || ||| | || ||||||||||| ||| |||||| |||| ||||||||||||||||||||
Db        661 CAAAAGGGCCAATGGAAGTGAACCAGGGGGTTACCAAGGGTTAAATGAAGAAGTAATAAC 720

Qy        660 AGG--CTCTGGAAAGAATTCTTGGAA 683
               ||      | ||||||||||||| |
Db        721 CGGGCTCTGGNAAAGAATTCTTGGGA 746


RESULT 6
BG719449
LOCUS       BG719449                 747 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602690172F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822528 5',
            mRNA sequence.
ACCESSION   BG719449
VERSION     BG719449.1  GI:13998636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 747)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10730  row: p  column: 17
            High quality sequence stop: 734.
FEATURES             Location/Qualifiers
     source          1..747
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4822528"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      220 a    121 c    226 g    180 t
ORIGIN

  Query Match             28.8%;  Score 624.8;  DB 12;  Length 747;
  Best Local Similarity   99.4%;  Pred. No. 4.2e-165;
  Matches  648;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    2;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 156

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 216

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 276

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 336

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 396

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 456

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAA-ATAATTCAGA 419
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        457 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAACATAATTCAGA 516

Qy        420 AGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATT 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 AGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATT 576

Qy        480 TGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGG 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 TGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGG 636

Qy        540 CCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAG 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 CCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAG 696

Qy        600 CAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAA 651
              |||||||||||||||||||||||||||||||||| ||||||||| |||||||
Db        697 CAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTAC-AAGGTTTAATTGAAGAA 747


RESULT 7
BG718234
LOCUS       BG718234                 682 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602696209F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4828298 5',
            mRNA sequence.
```

```
ACCESSION   BG718234
VERSION     BG718234.1  GI:13997421
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10746  row: a  column: 03
            High quality sequence stop: 682.
FEATURES             Location/Qualifiers
     source          1..682
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4828298"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      205 a    109 c    204 g    164 t
ORIGIN

  Query Match             28.6%;  Score 620.4;  DB 12;  Length 682;
  Best Local Similarity   99.8%;  Pred. No. 7.1e-164;
  Matches  621;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 120

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 180

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 240

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 300

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 360

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 420

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 480

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 540

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 600

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 660

Qy        601 AGAAGTGGCAGTGGAAGTGAAC 622
              ||||||||||||||||||||||
Db        661 AGAAGTGGCAGTGGAAGTGAAC 682


RESULT 8
BI560849
LOCUS       BI560849                 693 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603254018F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296444 5',
            mRNA sequence.
ACCESSION   BI560849
VERSION     BI560849.1  GI:15448163
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11750  row: c  column: 05
            High quality sequence stop: 691.
FEATURES             Location/Qualifiers
     source          1..693
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5296444"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      207 a    111 c    207 g    168 t
ORIGIN

  Query Match             28.6%;  Score 620.4;  DB 13;  Length 693;
  Best Local Similarity   99.8%;  Pred. No. 7.1e-164;
  Matches  621;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 131
```

```
Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 191

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 251

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 311

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 371

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 431

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 491

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 551

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 611

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 671

Qy        601 AGAAGTGGCAGTGGAAGTGAAC 622
              ||||||||||||||||||||||
Db        672 AGAAGTGGCAGTGGAAGTGAAC 693
```

RESULT 9
BI462207

```
LOCUS       BI462207                 830 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603205520F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5271221 5',
            mRNA sequence.
ACCESSION   BI462207
VERSION     BI462207.1  GI:15252863
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 830)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11684  row: h  column: 06
            High quality sequence stop: 769.
FEATURES             Location/Qualifiers
     source          1..830
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5271221"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      255 a    137 c     244 g     193 t      1 others
ORIGIN

  Query Match             28.3%;  Score 615.6;  DB 13;  Length 830;
  Best Local Similarity   95.8%;  Pred. No. 1.8e-162;
  Matches  729;  Conservative    0;  Mismatches   20;  Indels   12;  Gaps    9;

Qy          1 ATGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 131

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 191

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 251

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 311

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        312 ACAATGGGTGGTTNTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 371

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 431

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 491

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        492 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGA-TT 550

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 610

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGT-GGCACAGGTAATGGTGATACTTCTCAAAG 599
              | |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        611 C-TTTTGGTTCTAGAAGACCAGTATTAAGTGGGCACAGGTAATGGTGATACTTCTCAAAG 669

Qy        600 CAGAAGTGGCAG-TGGAAGTGAACG-AGGTGGTTACAAAGGTTT---AAATGAAGAAGTA 654
              |||||||||||| |||||||||||| |||||||||| || ||     | ||||||
Db        670 CAGAAGTGGCAGTTGGAAGTGAACGAAGGTGGTTACCAAAGGTTTCAACTGAACGAAGTA 729

Qy        655 ATAACAGGCTCTGGAA-AGAATTCTTGGAAGTCAGAA-GCAGAAGGAGGA--GAAAGTAG 710
              ||||||| |||||| ||| ||||||||||||||| |||||||||||    | |||||
Db        730 ATAACAGGGTCTGGAACAGACTTCTTGGAAGTCAGAAGGCAGAAGGAGGACGAACAGTAG 789

Qy        711 TGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTC 751
              |||||||||||||| ||||||| ||  ||||||| || ||
Db        790 TGATACTCAAGGACCCAAAGTGAACTTAATACCCCGTCTTC 830
```

RESULT 10

```
BI561001
LOCUS       BI561001                 918 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603254391F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296534 5',
            mRNA sequence.
ACCESSION   BI561001
VERSION     BI561001.1  GI:15448315
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 918)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11750  row: f  column: 23
            High quality sequence stop: 745.
FEATURES             Location/Qualifiers
     source          1..918
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5296534"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      288 a    160 c    253 g    217 t
ORIGIN

  Query Match             28.2%;  Score 612.4;  DB 13;  Length 918;
  Best Local Similarity   86.8%;  Pred. No. 1.5e-161;
  Matches  753;  Conservative    0;  Mismatches   36;  Indels   79;  Gaps    4;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 120

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 180

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 240

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||
Db        241 TCTGGGCGGAATTTTGGAAACAGAGATGCT------------------------------ 270

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
                                                         |||||||||||||||||
Db        271 -------------------------------------------GGTTTTCAAACAGCAGG 288

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 348

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 408

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 468

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 528

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 588

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
              |||||||||||||||||||||||||||||||||||||||||||     |  |  | |||
Db        589 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAACATGAAGAAGTAATACA 648

Qy        661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAG-AAAGTAGTGATACTCA 719
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        649 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAAGTAGTGATACTCA 708

Qy        720 AGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGG-----ACTCCATC 774
              |||||||||||||||||||||||||||||||||||   |||||| |      |   |
Db        709 AGGACCAAAAGTGACCTACATACCCCCTCCTCCACTGAAGGATGAAGGGACTCCAATCTT 768

Qy        775 TTTGCACATTATCAGACAGGCATAAAC-TTCGACAAATACGACACTATTCTTGTGGAAGT 833
              || || |||||||||||||||||||| |||||||||||||||||||||||||  |||||
Db        769 TTGGCCCATTATCAGACAGGCATAAACTTTCGACAAATACGACACTATTCTTGGTGAAGT 828

Qy        834 GTCTGGACATGATGCACCACCAGCAATT 861
              ||||||   |  |  ||||| |   ||
Db        829 GTCTGGAATTTAATGACCACAACAATTT 856


RESULT 11
BG718634
LOCUS       BG718634                 846 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602696739F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4828977 5',
            mRNA sequence.
ACCESSION   BG718634
VERSION     BG718634.1  GI:13997821
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10747  row: m  column: 10
            High quality sequence stop: 701.
FEATURES             Location/Qualifiers
     source          1..846
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4828977"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
```

```
                  /note="Organ: testis; Vector: pBluescriptR (modified
                  pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                  ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                  size-selected for  average insert size 2.2 kb and
                  normalized to ROT 5. This is a primary library enriched
                  for full-length clones and constructed using the
                  Cap-trapper method (Carninci, in preparation).  Library
                  constructed by M. Brownstein (NIMH/NHGRI, National
                  Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      251 a    153 c    248 g    194 t
ORIGIN

  Query Match             28.1%;  Score 610.4;  DB 12;  Length 846;
  Best Local Similarity   94.7%;  Pred. No. 5.3e-161;
  Matches  720;  Conservative    0;  Mismatches   26;  Indels   14;  Gaps    8;

Qy          1 ATGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 131

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 191

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 251

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 311

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 371

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 431

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 491

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGG-AAGAGGTAGTTTCCGAGGTTGCCGTGG-AGGAT 478
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||| |||||
Db        492 GCTTCAGGGCCATACAGAAGAGGTGGCAAGAGGTAGTTTCCGAGGTTGCCGTGGCAGGAT 551

Qy        479 TTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTG 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 TTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTG 611

Qy        539 GCCTTTTTGGTTCTAGAAGACCAGTATT-AAGTGGCACAGGTAATGGTGATACTTCTCAA 597
              ||||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        612 GCCTTTTGGGTTCTAGAAGACCAGTATTCAAGTGGCACAGGTAATGGTGATACTTCTCAA 671

Qy        598 AGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACA-AAGGTTTAAATGAAGAAGTAAT 656
              ||||||||||||||||||||||||||||||||||||| |||||||| |||| |||||| |
Db        672 AGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAGAAGGTTTACATGACGAAGTATT 731

Qy        657 AA-CAGGCTCTGGAAAG---AATTCTTGGAAGTCAGAAGCAGAAGGAGGAG----AAAGT 708
              || ||||||||||||||   | |  | |  |||||||||  | | | |  |    |  ||
Db        732 AACCAGGCTCTGGAAAGACAATTCTTGGACAGTCAGAAGGCGGACGCGACGGAGCACCGT 791

Qy        709 AGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTC 748
              || ||||||||  |  | | |||||||||||||||| |||
Db        792 AGCGATACTCA--GGACCACAGTGACCTACATACCCGCTC 829


RESULT 12
BG771822
LOCUS       BG771822                 699 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602720504F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837296 5',
            mRNA sequence.
ACCESSION   BG771822
VERSION     BG771822.1  GI:14082475
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10769  row: h  column: 01
            High quality sequence stop: 692.
FEATURES             Location/Qualifiers
     source          1..699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4837296"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      208 a    113 c    211 g    167 t
ORIGIN

  Query Match             27.5%;  Score 597.4;  DB 12;  Length 699;
  Best Local Similarity   99.5%;  Pred. No. 2.3e-157;
  Matches  620;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 ATGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 137

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 197

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 257

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 317

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 ACAATGGGTGG-TTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 376

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 436

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        437 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 496

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 556

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        557 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGG- 615

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 675

Qy        601 AGAAGTGGCAGTGGAAGTGAACG 623
              |||||||||||||||||||||||
Db        676 AGAAGTGGCAGTGGAAGTGAACG 698


RESULT 13
BI559479
LOCUS       BI559479                 758 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603252992F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5295407 5',
            mRNA sequence.
ACCESSION   BI559479
VERSION     BI559479.1  GI:15446793
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL).
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11747  row: g  column: 24
            High quality sequence stop: 741.
FEATURES             Location/Qualifiers
     source          1..758
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5295407"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      232 a    127 c    220 g    179 t
ORIGIN

  Query Match            27.5%;  Score 596.6;  DB 13;  Length 758;
  Best Local Similarity  91.2%;  Pred. No. 4e-157;
  Matches  681;  Conservative    0;  Mismatches    4;  Indels   62;  Gaps    2;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 121

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 181

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 241

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 301

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 361

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              |||||||||||||||||||||||||||||||||
Db        362 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGA--------------------------- 394

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
                                               |||||||||||||||||||||||||||
Db        395 ---------------------------------GGCTATCGAGATGGAAATAATTCAGAA 421

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 481

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 541

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 601

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAA--GAAGTAATAA 658
              |||||||||||||||||||||||||||||||||||||||||| |||||   |||||||||
Db        602 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTACATGAACGCAAGTAATAA 661

Qy        659 CAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTC 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTC 721

Qy        719 AAGGACCAAAAGTGACCTACATACCCC 745
              |||||||||||||||||||||||||||
Db        722 CAGGACCAAAAGTGACCTACATACCCC 748


RESULT 14
BI461393
LOCUS       BI461393                 861 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603206861F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272430 5',
            mRNA sequence.
ACCESSION   BI461393
VERSION     BI461393.1  GI:15252049
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 861)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
```

```
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11687  row: j  column: 15
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1..861
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5272430"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      265 a    147 c    247 g    201 t      1 others
ORIGIN

  Query Match             27.4%;  Score 594.8;  DB 13;  Length 861;
  Best Local Similarity   89.7%;  Pred. No. 1.4e-156;
  Matches  722;  Conservative    0;  Mismatches   63;  Indels   20;  Gaps    7;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 108

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 168

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 228

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 288

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        289 ACAATGGGTGGTTNTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 348

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 408

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 468

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        469 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGA-TT 527

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 587

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 647

Qy        601 AGAAGTGG--CAGTGGAAGTGAACGA--GGTGGTTACAAAGGTTTAAA----TGAAGAAG 652
              ||||||||       | |||||||||   |   ||||  | ||||||||     |  |||
Db        648 AGAAGTGGGCAGTGGAAAGTGAACGACGGTGGGTTCCCAAGGTTTACACTCGAAAGAAAG 707

Qy        653 TAATAACAGGCTCTGG--AAAGAATTCTTGG------AAGTCAGAAGCAGAAGGAGGAGA 704
              |||||| | || ||  ||||||||||||        ||      |   || |||  |  |
Db        708 TAATAACCAGGTCCGGGAAAAGAATTCTTGGGAAGTCCAGAAGCAGAAAGGAGGCAGCAA 767

Qy        705 AAGTAGTGATACTCAAGGACCAAAAGTGACC---TACATACCCCCTCCTCCACCTGAGGA 761
              | ||| ||||||||||      |||||  |     |  ||||||  | |||||   | ||
Db        768 ACGTATGGATACTCAAGGGGACCAAGTGGACCTTAATTTACCCCTGCTTCCACTGTAAGA 827

Qy        762 TGAGGACTCCATCTTTGCACATTAT 786
              ||||||    |   ||||  || | |
Db        828 TGAGGAACTCCACTTTTGGCAATTT 852


RESULT 15
BG717812
LOCUS       BG717812                 657 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602693829F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4826190 5',
            mRNA sequence.
ACCESSION   BG717812
VERSION     BG717812.1  GI:13996999
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10740  row: i  column: 07
            High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1..657
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4826190"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      200 a    105 c    192 g    160 t
ORIGIN

  Query Match             27.3%;  Score 593.4;  DB 12;  Length 657;
  Best Local Similarity   99.7%;  Pred. No. 3e-156;
  Matches  605;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 110

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 170

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        171 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 230

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 290

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 350

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 410

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCA-GA 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        411 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCACGA 470

Qy        420 AGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATT 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATT 530

Qy        480 TGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGG 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 TGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGG 590

Qy        540 CCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAG 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 CCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAG 650

Qy        600 CAGAAGT 606
              |||||||
Db        651 CAGAAGT 657
```

Search completed: June 10, 2003, 10:40:19
Job time : 2779 secs

```
> O <
O| |O  IntelliGenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Wed 11 Jun 103 10:51:36-PDT

Solution Parameters:

    Nucleic Alphabet   = Identity
    Output line length = 80
    Compress           = Off
    Histogram          = Off
    Randomization      = Off

    AMINO-Res-length   = 2
    DELetion-weight    = 5.00
    LEngth-factor      = 0
    Matching-weight    = 1.00
    NUCLEIC-Res-length = 4
    SPread-factor      = 50

Clustered order of selected sequences:

     1. US-09-714-865-1       (1-2224)
     7. US-09-714-865-15      (1-2172)

Region Alignment: (listed in Clustered order)

US-09-714-     1 acttgaagtcaccATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCC
                              |||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-     1              ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCC

 consensus       acttgaagtcaccATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCC

US-09-714-    62 TATGTTCCCATATTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-    49 TATGTTCCCATATTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGA

 consensus       TATGTTCCCATATTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGA

US-09-714-   123 CTCCAGCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAG
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   110 CTCCAGCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAG

 consensus       CTCCAGCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAG

US-09-714-   184 TGGATTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGAT
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   171 TGGATTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGAT

 consensus       TGGATTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGAT

US-09-714-   245 AATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   232 AATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAA

 consensus       AATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAA

US-09-714-   306 ACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   293 ACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGA

 consensus       ACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGA

US-09-714-   367 AGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   354 AGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT

 consensus       AGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT

US-09-714-   428 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAG
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAG

 consensus       TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAG

US-09-714-   489 GATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGG
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   476 GATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGG

 consensus       GATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGG

US-09-714-   550 TGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   537 TGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAA

 consensus       TGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAA

US-09-714-   611 AGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   598 AGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAA

 consensus       AGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAA

US-09-714-   672 CAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   659 CAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCA

 consensus       CAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCA

US-09-714-   733 AGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   720 AGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA

 consensus       AGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA

US-09-714-   794 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGAC

 consensus       CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGAC

US-09-714-   855 ATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAA
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   842 ATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAA

 consensus       ATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAA

US-09-714-   916 CAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714-   903 CAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATC

 consensus       CAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATC

US-09-714-   977 ATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTT
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
US-09-714-  964 ATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTT

 consensus      ATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTT

US-09-714- 1038 TTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1025 TTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGA

 consensus      TTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGA

US-09-714- 1099 GTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1086 GTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT

 consensus      GTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT

US-09-714- 1160 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAA

 consensus      TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAA

US-09-714- 1221 CCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1208 CCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCC

 consensus      CCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCC

US-09-714- 1282 TGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1269 TGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA

 consensus      TGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA

US-09-714- 1343 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1330 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAA

 consensus      GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAA

US-09-714- 1404 TTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1391 TTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTT

 consensus      TTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTT

US-09-714- 1465 TCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1452 TCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCT

 consensus      TCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCT

US-09-714- 1526 GTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1513 GTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGT

 consensus      GTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGT

US-09-714- 1587 TCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1574 TCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGT

 consensus      TCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGT

US-09-714- 1648 CTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1635 CTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATA

 consensus      CTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATA

US-09-714- 1709 TCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1696 TCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATT

 consensus      TCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATT

US-09-714- 1770 TTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1757 TTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATAT

 consensus      TTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATAT

US-09-714- 1831 TGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1818 TGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCAT

 consensus      TGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCAT

US-09-714- 1892 CGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1879 CGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTG

 consensus      CGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTG

US-09-714- 1953 AATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 1940 AATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGT

 consensus      AATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGT

US-09-714- 2014 TCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 2001 TCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACA

 consensus      TCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACA

US-09-714- 2075 AGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 2062 AGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG

 consensus      AGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG

US-09-714- 2136 GGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGATtaaagccaaaa
                ||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-714- 2123 GGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT

 consensus      GGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGATtaaagccaaaa

US-09-714- 2197 catccttcaagtctgtggttttgatgca

US-09-714- 2173

 consensus      catccttcaagtctgtggttttgatgca

Alignment score = 1912.00
```

Scoring matrix:

	1	7
1		1912
7		

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-714-1.res made by bshears on Wed 11 Jun 103 10:54:15-PDT.


Query sequence being compared:   US-714-1 (1-2224)
Number of sequences searched:                   1
Number of scores above cutoff:                  1

      Results of the initial comparison of US-714-1 (1-2224) with:
        File : /home/bshears/canella15.seq

   100-
      -
N     -
U   50-
M     -
B     -
E     -
R     -
      -
O     -
F   10-
      -
S     -
E    5-
Q     -
U     -
E     -
N     -
C     -
E     -                                                                *
S    0-------------------------------------------------------------------------
        |     |      |      |      |      |      |      |      |      |
SCORE 0    241    483    724    965    1207   1448   1689   1931   2172
STDEV


                              PARAMETERS

Similarity matrix       Unitary    K-tuple                    4
Mismatch penalty              1    Joining penalty           30
Gap penalty                1.00    Window size               32
Gap size penalty           0.33
Cutoff score                  0
Randomization group           0

                              SEARCH STATISTICS

Scores:                    Mean        Median      Standard Deviation
                           2172           0           0.00

Times:                      CPU                     Total Elapsed
                        00:00:00.00                 00:00:00.00

Number of residues:                  2172
Number of sequences searched:           1
Number of scores above cutoff:          1

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.


The list of best scores is:

                                                                  Init.  Opt.
Sequence Name     Description                              Length Score Score  Sig. Frame
------------------------------------------------------------------------------------------
  1. US-714-15     Sequence 15, Application US/0    2172  2172  2172   0.00    0


1. US-714-1 (1-2224)
   US-714-15    Sequence 15, Application US/09714865

 Sequence 15, Application US/09714865
 GENERAL INFORMATION:
  APPLICANT:  Castrillon, Diego H.
  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
  TITLE OF INVENTION:  IMPROVED DIAGNOSIS OF GERM CELL TUMORS
  FILE REFERENCE:  B0801/7195/ERP/KA
  CURRENT APPLICATION NUMBER: US/09/714,865
  CURRENT FILING DATE:  2000-11-16
  PRIOR APPLICATION NUMBER:  U.S. 60/166,394
  PRIOR FILING DATE:  1999-11-18
  NUMBER OF SEQ ID NOS:  15
  SOFTWARE:  FastSEQ for Windows Version 3.0
 SEQ ID NO 15
   LENGTH: 2172
   TYPE: DNA
   ORGANISM: Homo Sapiens

Initial Score    =    2172  Optimized Score  =    2172  Significance =  0.00
Residue Identity =    100%  Matches          =    2172  Mismatches   =     0
Gaps             =       0  Conservative Substitutions              =     0

            10   X     20        30        40        50        60        70
    ACTTGAAGTCACCATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCAT
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCAT
                 X        10        20        30        40        50

          80        90       100       110       120       130       140
    ATTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAAT
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    ATTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAAT
  60        70        80        90       100       110       120       130

       150       160       170       180       190       200       210
    GGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGAAACAG
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    GGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGAAACAG
         140       150       160       170       180       190       200

     220       230       240       250       260       270       280
    AGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGG
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    AGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGG
       210       220       230       240       250       260       270

   290       300       310       320       330       340       350       360
    AAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGA
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    AAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGA
      280       290       300       310       320       330       340

          370       380       390       400       410       420       430
    CTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGA
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    CTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGA
    350       360       370       380       390       400       410

        440       450       460       470       480       490       500
    AGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAG
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    AGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAG
  420       430       440       450       460       470       480       490

       510       520       530       540       550       560       570
```

```
TCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGT
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGT
      500       510       520       530       540       550       560

 580       590       600       610       620       630       640
ATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAA
    570       580       590       600       610       620       630

650       660       670       680       690       700       710       720
AGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAG
  640       650       660       670       680       690       700

       730       740       750       760       770       780       790
TAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGC
710       720       730       740       750       760       770

     800       810       820       830       840       850       860
ACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACC
780       790       800       810       820       830       840       850

   870       880       890       900       910       920       930
ACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTA
      860       870       880       890       900       910       920

 940       950       960       970       980       990       1000
TACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGC
    930       940       950       960       970       980       990

1010      1020      1030      1040      1050      1060      1070      1080
TCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAAC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAAC
  1000      1010      1020      1030      1040      1050      1060

       1090      1100      1110      1120      1130      1140      1150
TGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCA
1070      1080      1090      1100      1110      1120      1130

       1160      1170      1180      1190      1200      1210      1220
GATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCA
1140      1150      1160      1170      1180      1190      1200      1210

  1230      1240      1250      1260      1270      1280      1290
GCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGA
      1220      1230      1240      1250      1260      1270      1280

1300      1310      1320      1330      1340      1350      1360
TATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTT
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTT
   1290      1300      1310      1320      1330      1340      1350

1370      1380      1390      1400      1410      1420      1430      1440
GGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCA
  1360      1370      1380      1390      1400      1410      1420

      1450      1460      1470      1480      1490      1500      1510
AACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTA
1430      1440      1450      1460      1470      1480      1490

    1520      1530      1540      1550      1560      1570      1580
TCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCA
1500      1510      1520      1530      1540      1550      1560      1570

  1590      1600      1610      1620      1630      1640      1650
GTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGA
     1580      1590      1600      1610      1620      1630      1640

1660      1670      1680      1690      1700      1710      1720
AACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGG
   1650      1660      1670      1680      1690      1700      1710

1730      1740      1750      1760      1770      1780      1790      1800
TGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGC
  1720      1730      1740      1750      1760      1770      1780

      1810      1820      1830      1840      1850      1860      1870
TACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTAC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTAC
1790      1800      1810      1820      1830      1840      1850

     1880      1890      1900      1910      1920      1930      1940
CATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTT
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTT
1860      1870      1880      1890      1900      1910      1920      1930

  1950      1960      1970      1980      1990      2000      2010
TGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCC
     1940      1950      1960      1970      1980      1990      2000

2020      2030      2040      2050      2060      2070      2080
TGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTT
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTT
   2010      2020      2030      2040      2050      2060      2070

2090      2100      2110      2120      2130      2140      2150      2160
TGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAA
  2080      2090      2100      2110      2120      2130      2140

      2170      2180    X 2190      2200      2210      2220
TCCAGTAGATGATGAGTCATGGGATTAAAGCCAAAACATCCTTCAAGTCTGTGGTTTTGATGCA
|||||||||||||||||||||||||
TCCAGTAGATGATGAGTCATGGGAT
2150      2160      2170 X
```

```
                     GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protèin search, using frame_plus_n2p model

Run on:         June 10, 2003, 16:54:21 ; Search time 74 Seconds
                                          (without alignments)
                                          12095.523 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  724
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  OLIGO
                Xgapop 60.0 , Xgapext 60.0
                Ygapop 60.0 , Ygapext 60.0
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       671580 seqs, 206047115 residues

Word size:      1

Total number of hits satisfying chosen parameters:      1343044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111833_26583/app_query.fasta_1.2311
-DB=SPTREMBL_21 -QFMT=fastan -SUFFIX=oligo.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=1000 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_119_@runat_05062003_111833_26583 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :      SPTREMBL_21:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	3.7	546	5	Q9GV10	Q9gv10 ephydatia f
2	25	3.5	394	13	Q8QGD0	Q8qgd0 pantodon bu
3	24	3.3	399	13	Q9I8L8	Q9i8l8 danio dangi
4	24	3.3	617	13	Q90ZF6	Q90zf6 oryzias lat
5	24	3.3	621	13	Q8QHL6	Q8qhl6 oreochromis
6	24	3.3	645	13	Q9DEG3	Q9deg3 oreochromis
7	24	3.3	647	13	Q9PT10	Q9pt10 oncorhynchu
8	24	3.3	700	13	Q91372	Q91372 xenopus lae
9	24	3.3	700	13	O42378	O42378 brachydanio
10	24	3.3	715	13	Q8QGG8	Q8qgg8 brachydanio
11	24	3.3	715	13	Q8QFU0	Q8qfu0 brachydanio
12	24	3.3	716	13	O42107	O42107 brachydanio
13	23	3.2	400	13	Q8QGC9	Q8qgc9 melanotaeni
14	22	3.0	658	11	Q9QWS9	Q9qws9 mus musculu
15	21	2.9	396	13	Q8QGC8	Q8qgc8 oncorhynchu
16	20	2.8	659	5	O96069	O96069 ciona intes
17	20	2.8	669	5	O96068	O96068 ciona intes
18	20	2.8	688	5	Q9GNP2	Q9gnp2 ciona savig
19	20	2.8	688	13	O42375	O42375 brachydanio
20	20	2.8	770	5	Q9GNP1	Q9gnp1 ciona savig
21	19	2.6	376	5	Q8SXU8	Q8sxu8 drosophila
22	19	2.6	396	13	Q8QGD1	Q8qgd1 oryzias lat
23	19	2.6	397	13	Q8QGD2	Q8qgd2 hyphessobry
24	19	2.6	398	13	Q8QGD3	Q8qgd3 cyprinus ca
25	19	2.6	573	5	Q9GV07	Q9gv07 dugesia dor
26	19	2.6	646	10	Q9M2F9	Q9m2f9 arabidopsis
27	19	2.6	798	5	Q9VHP0	Q9vhp0 drosophila
28	19	2.6	798	5	Q8SXI8	Q8sxi8 drosophila
29	18	2.5	491	5	Q9GV11	Q9gv11 ephydatia f
30	18	2.5	628	5	Q9GV14	Q9gv14 hydra magni
31	18	2.5	633	10	Q9SIN6	Q9sin6 arabidopsis
32	18	2.5	726	5	O97031	O97031 dugesia jap
33	18	2.5	763	5	Q22873	Q22873 caenorhabdi
34	18	2.5	974	5	Q966L9	Q966l9 caenorhabdi
35	18	2.5	974	5	Q27376	Q27376 caenorhabdi
36	17	2.3	603	10	Q9M2G5	Q9m2g5 arabidopsis
37	17	2.3	662	13	Q9DGR8	Q9dgr8 gallus gall
38	17	2.3	688	3	Q9P6U9	Q9p6u9 neurospora
39	16	2.2	225	3	Q9UU14	Q9uu14 schizosacch
40	16	2.2	617	3	Q8TFK8	Q8tfk8 candida gla
41	16	2.2	641	5	Q9N3F4	Q9n3f4 caenorhabdi
42	16	2.2	644	5	Q95XY1	Q95xy1 caenorhabdi
43	15	2.1	108	5	Q9NBW6	Q9nbw6 drosophila
44	15	2.1	255	3	O93932	O93932 candida gla
45	15	2.1	263	10	Q9ZRZ7	Q9zrz7 arabidopsis
46	15	2.1	328	5	Q9BL38	Q9bl38 caenorhabdi
47	15	2.1	480	10	Q8S0B0	Q8s0b0 oryza sativ
48	15	2.1	484	10	Q8W4R3	Q8w4r3 arabidopsis
49	15	2.1	494	10	Q94IW4	Q94iw4 oryza sativ
50	15	2.1	501	10	Q9C718	Q9c718 arabidopsis
51	15	2.1	561	5	Q9XUW5	Q9xuw5 caenorhabdi
52	15	2.1	564	10	Q9FML1	Q9fml1 arabidopsis
53	15	2.1	578	5	Q9I7P5	Q9i7p5 drosophila
54	15	2.1	595	13	Q9W744	Q9w744 gallus gall
55	15	2.1	608	13	Q9I8S3	Q9i8s3 xenopus lae
56	15	2.1	618	5	Q9XTP2	Q9xtp2 molgula ocu
57	15	2.1	618	10	Q93WJ3	Q93wj3 arabidopsis
58	15	2.1	619	5	Q9Y1M2	Q9y1m2 molgula occ
59	15	2.1	619	10	O64430	O64430 arabidopsis
60	15	2.1	619	10	Q8W4F5	Q8w4f5 arabidopsis
61	15	2.1	630	5	Q9N5K1	Q9n5k1 caenorhabdi
62	15	2.1	703	5	Q9VHV7	Q9vhv7 drosophila
63	15	2.1	713	10	Q9LYJ9	Q9lyj9 arabidopsis
64	15	2.1	716	10	Q9SP26	Q9sp26 pisum sativ
65	15	2.1	733	10	Q942R9	Q942r9 oryza sativ
66	15	2.1	781	5	O97032	O97032 dugesia jap
67	15	2.1	811	5	Q9VRZ5	Q9vrz5 drosophila
68	15	2.1	891	10	Q8RYT6	Q8ryt6 oryza sativ
69	15	2.1	945	5	Q9W3M7	Q9w3m7 drosophila
70	14	1.9	421	2	Q93AA5	Q93aa5 yersinia pe
71	14	1.9	421	10	Q8W4P3	Q8w4p3 arabidopsis
72	14	1.9	422	16	Q9KLH6	Q9klh6 vibrio chol
73	14	1.9	423	16	Q9KU63	Q9ku63 vibrio chol
74	14	1.9	425	2	Q9ZI52	Q9zi52 anabaena va

75	14	1.9	426	16	Q8YN52	Q8yn52 anabaena sp
76	14	1.9	446	16	Q9PGV6	Q9pgv6 xylella fas
77	14	1.9	446	16	Q9HYE0	Q9hye0 pseudomonas
78	14	1.9	451	16	Q9KMW4	Q9kmw4 vibrio chol
79	14	1.9	454	16	Q8ZQN8	Q8zqn8 salmonella
80	14	1.9	455	16	Q8Z877	Q8z877 salmonella
81	14	1.9	455	16	Q8X7Y7	Q8x7y7 escherichia
82	14	1.9	458	16	Q9I663	Q9i663 pseudomonas
83	14	1.9	460	16	Q8ZFD8	Q8zfd8 yersinia pe
84	14	1.9	462	16	Q9JYY7	Q9jyy7 neisseria m
85	14	1.9	462	16	Q9JTY3	Q9jty3 neisseria m
86	14	1.9	465	16	Q8Y1R7	Q8y1r7 ralstonia s
87	14	1.9	495	5	Q8SRB2	Q8srb2 encephalito
88	14	1.9	495	16	Q8XXA5	Q8xxa5 ralstonia s
89	14	1.9	522	1	Q9P9G7	Q9p9g7 methanococc
90	14	1.9	535	16	Q8YH70	Q8yh70 brucella me
91	14	1.9	540	16	Q8Y1Z8	Q8y1z8 ralstonia s
92	14	1.9	639	16	Q9I689	Q9i689 pseudomonas
93	14	1.9	732	5	Q9N478	Q9n478 caenorhabdi
94	14	1.9	1088	10	Q9SQV1	Q9sqv1 arabidopsis
95	13	1.8	368	17	O28030	O28030 archaeoglob
96	13	1.8	397	16	Q9KS53	Q9ks53 vibrio chol
97	13	1.8	397	16	Q9HXE5	Q9hxe5 pseudomonas
98	13	1.8	411	10	Q9ZS13	Q9zs13 arabidopsis
99	13	1.8	457	16	Q9JYU9	Q9jyu9 neisseria m
100	13	1.8	457	16	Q9JTU1	Q9jtu1 neisseria m
101	13	1.8	463	5	Q9BL37	Q9bl37 caenorhabdi
102	13	1.8	474	10	Q9SWV9	Q9swv9 lycopersico
103	13	1.8	484	5	Q95XD7	Q95xd7 caenorhabdi
104	13	1.8	504	5	Q9XVZ6	Q9xvz6 caenorhabdi
105	13	1.8	537	10	Q9C551	Q9c551 arabidopsis
106	13	1.8	601	5	O01378	O01378 bombyx mori
107	13	1.8	622	10	Q9LKL6	Q9lkl6 pisum sativ
108	13	1.8	747	5	Q965K1	Q965k1 caenorhabdi
109	13	1.8	797	5	Q9GV13	Q9gv13 hydra magni
110	13	1.8	970	5	Q965K2	Q965k2 caenorhabdi
111	12	1.7	202	2	Q93S73	Q93s73 xanthomonas
112	12	1.7	206	2	P74957	P74957 shewanella
113	12	1.7	337	11	Q9CSD1	Q9csd1 mus musculu
114	12	1.7	386	5	Q9U936	Q9u936 plasmodium
115	12	1.7	411	2	O87951	O87951 vibrio algi
116	12	1.7	441	16	Q8ZD82	Q8zd82 yersinia pe
117	12	1.7	444	16	Q8ZMX7	Q8zmx7 salmonella
118	12	1.7	444	16	Q8Z4J8	Q8z4j8 salmonella
119	12	1.7	444	16	Q8XA21	Q8xa21 escherichia
120	12	1.7	446	2	Q8RTR7	Q8rtr7 uncultured
121	12	1.7	542	10	Q9SZB4	Q9szb4 arabidopsis
122	12	1.7	580	4	Q9H0W3	Q9h0w3 homo sapien
123	12	1.7	591	10	Q9LU46	Q9lu46 arabidopsis
124	12	1.7	619	11	Q9DBN9	Q9dbn9 mus musculu
125	12	1.7	626	11	Q8R0R6	Q8r0r6 mus musculu
126	12	1.7	730	5	Q95QN2	Q95qn2 caenorhabdi
127	12	1.7	769	5	Q95QN3	Q95qn3 caenorhabdi
128	12	1.7	844	5	Q9NNF5	Q9nnf5 leishmania
129	12	1.7	890	5	Q9GV12	Q9gv12 hydra magni
130	12	1.7	1031	4	Q9Y658	Q9y658 homo sapien
131	12	1.7	1031	4	Q96EI0	Q96ei0 homo sapien
132	12	1.7	1032	4	O94894	O94894 homo sapien
133	12	1.7	1032	11	Q62780	Q62780 rattus norv
134	11	1.5	121	5	Q9GV01	Q9gv01 aurelia aur
135	11	1.5	122	5	Q9GUZ6	Q9guz6 sanderia ma
136	11	1.5	343	17	Q97X64	Q97x64 sulfolobus
137	11	1.5	360	17	Q97WT0	Q97wt0 sulfolobus
138	11	1.5	449	16	Q9HX66	Q9hx66 pseudomonas
139	11	1.5	452	16	Q9KNA4	Q9kna4 vibrio chol
140	11	1.5	457	16	Q8ZP87	Q8zp87 salmonella
141	11	1.5	457	16	Q8X8Q4	Q8x8q4 escherichia
142	11	1.5	465	3	Q9P6N8	Q9p6n8 schizosacch
143	11	1.5	467	16	Q98N14	Q98n14 rhizobium l
144	11	1.5	482	16	Q8YGX0	Q8ygx0 brucella me
145	11	1.5	491	10	Q9SA27	Q9sa27 arabidopsis
146	11	1.5	499	16	Q92PK3	Q92pk3 rhizobium m
147	11	1.5	503	16	Q8UEC7	Q8uec7 agrobacteri
148	11	1.5	517	16	Q9A779	Q9a779 caulobacter
149	11	1.5	527	16	Q98NA7	Q98na7 rhizobium l
150	11	1.5	608	16	Q8Y0U1	Q8y0u1 ralstonia s
151	11	1.5	610	16	Q9CLU3	Q9clu3 pasteurella
152	11	1.5	646	16	Q8ZLT4	Q8zlt4 salmonella
153	11	1.5	646	16	Q8Z3I1	Q8z3i1 salmonella
154	11	1.5	646	16	Q8XA87	Q8xa87 escherichia
155	11	1.5	648	4	Q9NXZ2	Q9nxz2 homo sapien
156	11	1.5	662	3	Q9Y7T7	Q9y7t7 schizosacch
157	11	1.5	664	16	Q8ZBC8	Q8zbc8 yersinia pe
158	11	1.5	1016	5	Q9XTX5	Q9xtx5 caenorhabdi
159	11	1.5	1156	5	O44758	O44758 caenorhabdi
160	10	1.4	110	5	O01955	O01955 bombyx mori
161	10	1.4	111	2	Q9L6R9	Q9l6r9 salmonella
162	10	1.4	113	5	Q9GUZ4	Q9guz4 aurelia aur
163	10	1.4	113	5	Q9GUY7	Q9guy7 sanderia ma
164	10	1.4	114	5	Q9GUZ3	Q9guz3 acropora di
165	10	1.4	114	5	Q9GUZ2	Q9guz2 craspedacus
166	10	1.4	114	5	Q9GUZ0	Q9guz0 hydractinia
167	10	1.4	114	5	Q9GUY9	Q9guy9 tima formos
168	10	1.4	115	5	O01956	O01956 bombyx mori
169	10	1.4	123	5	O01958	O01958 bombyx mori
170	10	1.4	155	10	Q9C886	Q9c886 arabidopsis
171	10	1.4	187	10	Q9ZS02	Q9zs02 arabidopsis
172	10	1.4	224	10	Q9ZS09	Q9zs09 arabidopsis
173	10	1.4	240	4	Q9NT96	Q9nt96 homo sapien
174	10	1.4	376	10	Q9ZS10	Q9zs10 arabidopsis
175	10	1.4	421	16	Q8XAT4	Q8xat4 escherichia
176	10	1.4	423	16	Q9CJS1	Q9cjs1 pasteurella
177	10	1.4	425	17	O26305	O26305 methanobact
178	10	1.4	428	16	Q8ZAD8	Q8zad8 yersinia pe
179	10	1.4	438	16	Q9KV52	Q9kv52 vibrio chol
180	10	1.4	456	10	Q9FME1	Q9fme1 arabidopsis
181	10	1.4	460	16	Q9KP15	Q9kp15 vibrio chol
182	10	1.4	470	16	Q92CF8	Q92cf8 listeria in
183	10	1.4	470	16	Q8Y7M8	Q8y7m8 listeria mo
184	10	1.4	478	16	Q8XM28	Q8xm28 clostridium
185	10	1.4	497	2	Q9Z6C9	Q9z6c9 lactobacill
186	10	1.4	502	16	Q8UCZ5	Q8ucz5 agrobacteri
187	10	1.4	503	16	Q92UF7	Q92uf7 rhizobium m
188	10	1.4	511	16	P96614	P96614 bacillus su
189	10	1.4	513	16	Q8YXJ0	Q8yxj0 anabaena sp
190	10	1.4	516	16	Q92DF4	Q92df4 listeria in
191	10	1.4	520	16	Q8Y8N0	Q8y8n0 listeria mo
192	10	1.4	522	2	Q93HQ5	Q93hq5 streptococc
193	10	1.4	527	16	O05855	O05855 mycobacteri
194	10	1.4	528	16	Q97EU6	Q97eu6 clostridium
195	10	1.4	528	16	Q8XKJ8	Q8xkj8 clostridium
196	10	1.4	528	16	Q8RHN6	Q8rhn6 fusobacteri
197	10	1.4	539	16	Q9KAA6	Q9kaa6 bacillus ha
198	10	1.4	539	16	Q99Z38	Q99z38 streptococc
199	10	1.4	544	16	Q9CCH3	Q9cch3 mycobacteri
200	10	1.4	547	16	Q9CIK6	Q9cik6 lactococcus
201	10	1.4	555	17	Q8TJ32	Q8tj32 methanosarc
202	10	1.4	567	16	Q9I003	Q9i003 pseudomonas
203	10	1.4	578	3	Q9Y7C4	Q9y7c4 candida alb
204	10	1.4	589	10	Q9LGR4	Q9lgr4 oryza sativ
205	10	1.4	610	10	Q9FZ92	Q9fz92 arabidopsis
206	10	1.4	614	16	Q9PGP6	Q9pgp6 xylella fas
207	10	1.4	622	11	Q91VN6	Q91vn6 mus musculu
208	10	1.4	680	5	Q95TU3	Q95tu3 drosophila
209	10	1.4	760	10	O22907	O22907 arabidopsis
210	10	1.4	763	5	Q95XM9	Q95xm9 caenorhabdi
211	10	1.4	820	4	O43188	O43188 homo sapien
212	10	1.4	820	4	Q9BUQ8	Q9buq8 homo sapien
213	10	1.4	845	10	O49289	O49289 arabidopsis
214	10	1.4	855	5	P90549	P90549 leishmania
215	10	1.4	989	10	Q9SF41	Q9sf41 arabidopsis
216	10	1.4	1014	3	Q9P7C7	Q9p7c7 schizosacch
217	10	1.4	1166	10	Q9SYP6	Q9syp6 arabidopsis
218	10	1.4	1224	5	Q9VXW2	Q9vxw2 drosophila
219	9	1.2	49	2	Q9EYX4	Q9eyx4 salmonella
220	9	1.2	93	3	Q9USB9	Q9usb9 schizosacch

	221	9	1.2	106	5	O97029	O97029 dugesia jap
	222	9	1.2	107	5	O97026	O97026 dugesia jap
	223	9	1.2	108	5	O97021	O97021 dugesia jap
	224	9	1.2	108	5	O97030	O97030 dugesia jap
	225	9	1.2	110	5	O97022	O97022 dugesia jap
c	226	9	1.3	111	13	Q9YH66	Q9yh66 brachydanio
	227	9	1.2	114	5	O02028	O02028 bombyx mori
c	228	9	1.3	162	10	Q9M0B4	Q9m0b4 arabidopsis
	229	9	1.2	166	4	Q9BVQ8	Q9bvq8 homo sapien
	230	9	1.2	196	2	Q9EVK4	Q9evk4 rhizobium l
c	231	9	1.3	228	10	O81323	O81323 arabidopsis
	232	9	1.2	257	4	Q9NVY6	Q9nvy6 homo sapien
c	233	9	1.3	262	5	Q9VIX5	Q9vix5 drosophila
c	234	9	1.3	287	5	Q9VIX6	Q9vix6 drosophila
c	235	9	1.3	301	10	Q8S2F7	Q8s2f7 oryza sativ
	236	9	1.2	312	4	Q8WYG4	Q8wyg4 homo sapien
c	237	9	1.3	323	12	Q67714	Q67714 human adeno
c	238	9	1.3	323	12	Q67733	Q67733 human adeno
c	239	9	1.3	323	12	Q91CL7	Q91cl7 human adeno
	240	9	1.2	375	4	Q96F57	Q96f57 homo sapien
c	241	9	1.3	375	17	Q8TI63	Q8ti63 methanosarc
	242	9	1.2	376	16	O34750	O34750 bacillus su
c	243	9	1.3	378	5	O76736	O76736 dictyosteli
	244	9	1.2	406	4	Q9NWU3	Q9nwu3 homo sapien
c	245	9	1.3	406	5	Q9NG14	Q9ng14 branchiosto
c	246	9	1.3	412	3	Q9P603	Q9p603 neurospora
	247	9	1.2	441	16	Q8Y6G5	Q8y6g5 listeria mo
	248	9	1.2	442	16	Q9CJZ8	Q9cjz8 pasteurella
	249	9	1.2	442	16	Q92AT6	Q92at6 listeria in
c	250	9	1.3	448	5	Q9VYK1	Q9vyk1 drosophila
	251	9	1.2	453	3	Q9HGM5	Q9hgm5 schizosacch
c	252	9	1.3	467	11	Q64167	Q64167 mus musculu
	253	9	1.2	472	5	O96967	O96967 drosophila
	254	9	1.2	476	13	Q98TR4	Q98tr4 fugu rubrip
	255	9	1.2	481	16	Q97FI9	Q97fi9 clostridium
	256	9	1.2	483	4	Q9Y6V7	Q9y6v7 homo sapien
	257	9	1.2	486	5	Q9VKE2	Q9vke2 drosophila
	258	9	1.2	503	5	Q8SSG7	Q8ssg7 encephalito
c	259	9	1.3	508	11	Q9JJJ6	Q9jjj6 mus musculu
	260	9	1.2	517	3	Q8X1T4	Q8x1t4 emericella
c	261	9	1.3	543	16	Q8XW22	Q8xw22 ralstonia s
c	262	9	1.3	548	5	Q9W3D0	Q9w3d0 drosophila
c	263	9	1.3	556	10	Q9LGY7	Q9lgy7 oryza sativ
c	264	9	1.3	558	13	Q9PUB5	Q9pub5 brachydanio
	265	9	1.2	561	4	Q9BXF0	Q9bxf0 homo sapien
	266	9	1.2	572	11	Q921N6	Q921n6 mus musculu
	267	9	1.2	582	5	Q23909	Q23909 dictyosteli
c	268	9	1.3	606	16	Q8REH6	Q8reh6 fusobacteri
c	269	9	1.3	609	10	Q9ARZ1	Q9arz1 oryza sativ
	270	9	1.2	625	5	Q9NE37	Q9ne37 leishmania
c	271	9	1.3	684	5	Q8T848	Q8t848 dictyosteli
c	272	9	1.3	691	5	Q9NKR7	Q9nkr7 leishmania
	273	9	1.2	696	5	Q9N5M1	Q9n5m1 caenorhabdi
	274	9	1.2	709	4	O75619	O75619 homo sapien
	275	9	1.2	733	10	P93008	P93008 arabidopsis
	276	9	1.2	736	10	Q94GY4	Q94gy4 oryza sativ
	277	9	1.2	746	5	P91340	P91340 caenorhabdi
	278	9	1.2	758	10	Q8SAX7	Q8sax7 oryza sativ
	279	9	1.2	765	4	Q969N7	Q969n7 homo sapien
	280	9	1.2	765	4	Q96GQ7	Q96gq7 homo sapien
c	281	9	1.3	781	11	O89090	O89090 mus musculu
	282	9	1.2	782	5	Q9V378	Q9v378 drosophila
c	283	9	1.3	784	11	O89087	O89087 mus musculu
	284	9	1.2	789	10	Q9ZRZ8	Q9zrz8 arabidopsis
	285	9	1.2	791	5	Q9VTC1	Q9vtc1 drosophila
	286	9	1.2	796	4	Q96L97	Q96l97 homo sapien
c	287	9	1.3	809	10	Q43715	Q43715 pisum sativ
	288	9	1.2	811	5	Q93382	Q93382 caenorhabdi
c	289	9	1.3	812	10	Q8RYT4	Q8ryt4 oryza sativ
	290	9	1.2	819	4	Q96BK1	Q96bk1 homo sapien
	291	9	1.2	853	5	Q9N2X0	Q9n2x0 caenorhabdi
c	292	9	1.3	871	5	O44358	O44358 nephila cla
	293	9	1.2	899	5	Q95XS9	Q95xs9 caenorhabdi
c	294	9	1.3	902	5	Q8SX13	Q8sx13 drosophila
c	295	9	1.3	1152	5	Q95YY5	Q95yy5 leishmania
c	296	9	1.3	1884	5	Q9NHW2	Q9nhw2 nephila mad
c	297	9	1.3	2136	10	Q8RYW8	Q8ryw8 oryza sativ
c	298	9	1.3	2205	5	Q9NGQ2	Q9ngq2 dictyosteli
c	299	9	1.3	3112	5	Q9NKP1	Q9nkp1 leishmania
c	300	9	1.3	3469	5	Q9U4I2	Q9u4i2 drosophila
c	301	9	1.3	3502	5	Q9VYJ9	Q9vyj9 drosophila
	302	8	1.1	71	12	Q9YVY2	Q9yvy2 melanoplus
	303	8	1.1	78	2	Q54854	Q54854 streptococc
	304	8	1.1	92	4	Q9BVZ7	Q9bvz7 homo sapien
c	305	8	1.1	100	17	Q97WM9	Q97wm9 sulfolobus
	306	8	1.1	110	5	O97025	O97025 dugesia jap
	307	8	1.1	110	5	O97027	O97027 dugesia jap
	308	8	1.1	111	6	O77725	O77725 gorilla gor
	309	8	1.1	111	6	O77731	O77731 pan troglod
	310	8	1.1	111	6	O77703	O77703 canis famil
	311	8	1.1	111	11	O88757	O88757 rattus norv
	312	8	1.1	113	5	O01957	O01957 bombyx mori
	313	8	1.1	113	5	Q9GV08	Q9gv08 dugesia dor
	314	8	1.1	114	5	Q9GUZ1	Q9guz1 eirene sp.
	315	8	1.1	114	10	Q9MAW9	Q9maw9 oryza sativ
	316	8	1.1	114	10	Q9MAW8	Q9maw8 oryza sativ
	317	8	1.1	117	5	Q9GV09	Q9gv09 dugesia dor
	318	8	1.1	118	5	Q9GV00	Q9gv00 acropora di
	319	8	1.1	119	5	Q9GUZ7	Q9guz7 hydractinia
	320	8	1.1	120	5	Q9GUZ9	Q9guz9 craspedacus
	321	8	1.1	120	5	Q9GUZ8	Q9guz8 eirene sp.
	322	8	1.1	120	5	Q9GUZ5	Q9guz5 tima formos
	323	8	1.1	122	16	Q8YV51	Q8yv51 anabaena sp
c	324	8	1.1	126	4	Q16273	Q16273 homo sapien
c	325	8	1.1	129	10	O82295	O82295 arabidopsis
c	326	8	1.1	131	11	Q9DA59	Q9da59 mus musculu
	327	8	1.1	142	2	Q93SJ6	Q93sj6 myxococcus
c	328	8	1.1	145	5	Q9V722	Q9v722 drosophila
c	329	8	1.1	147	12	Q8VAS8	Q8vas8 white spot
c	330	8	1.1	159	5	Q9NAG0	Q9nag0 caenorhabdi
	331	8	1.1	169	11	O35662	O35662 mus musculu
c	332	8	1.1	174	5	O01544	O01544 caenorhabdi
	333	8	1.1	176	16	Q9RJ74	Q9rj74 streptomyce
c	334	8	1.1	177	4	Q96BW8	Q96bw8 homo sapien
	335	8	1.1	178	16	Q9RJQ5	Q9rjq5 streptomyce
c	336	8	1.1	188	4	O60377	O60377 homo sapien
	337	8	1.1	188	10	Q9FQ88	Q9fq88 nicotiana t
	338	8	1.1	191	10	Q9FQ83	Q9fq83 nicotiana t
	339	8	1.1	194	16	Q8RFY9	Q8rfy9 fusobacteri
c	340	8	1.1	195	11	Q9WUJ6	Q9wuj6 mus musculu
	341	8	1.1	214	5	Q9N7D9	Q9n7d9 leishmania
c	342	8	1.1	224	10	Q8W0N9	Q8w0n9 oryza sativ
c	343	8	1.1	226	4	Q9H4A8	Q9h4a8 homo sapien
	344	8	1.1	226	10	Q9FQ87	Q9fq87 nicotiana t
	345	8	1.1	229	15	Q9DKV7	Q9dkv7 caprine art
	346	8	1.1	242	10	Q94DI3	Q94di3 oryza sativ
	347	8	1.1	251	10	Q8W2W7	Q8w2w7 oryza sativ
	348	8	1.1	259	10	Q9ZS05	Q9zs05 arabidopsis
	349	8	1.1	263	8	Q9B8G8	Q9b8g8 heterodoxus
c	350	8	1.1	264	10	Q9FSL0	Q9fsl0 oryza sativ
	351	8	1.1	267	13	O57652	O57652 fugu rubrip
c	352	8	1.1	267	16	Q8Z073	Q8z073 anabaena sp
c	353	8	1.1	268	10	Q9SBF6	Q9sbf6 arabidopsis
c	354	8	1.1	273	11	Q9Z0V5	Q9z0v5 rattus norv
c	355	8	1.1	274	6	Q9BGI2	Q9bgi2 bos taurus
	356	8	1.1	274	16	P75019	P75019 mycobacteri
c	357	8	1.1	280	11	Q91VQ2	Q91vq2 mus musculu
c	358	8	1.1	283	5	Q9NB39	Q9nb39 plasmodium
	359	8	1.1	289	10	Q9FQ85	Q9fq85 nicotiana t
c	360	8	1.1	290	10	O65659	O65659 arabidopsis
	361	8	1.1	294	10	Q9FQ89	Q9fq89 nicotiana t
	362	8	1.1	296	10	Q9AUM8	Q9aum8 oryza sativ
c	363	8	1.1	296	11	Q99K76	Q99k76 mus musculu
c	364	8	1.1	296	11	Q9CXH8	Q9cxh8 mus musculu
c	365	8	1.1	296	11	Q64012	Q64012 mus musculu
c	366	8	1.1	299	2	Q9AMJ7	Q9amj7 pseudomonas

c	367	8	1.1	299	2	Q9AMJ6	Q9amj6 pseudomonas
c	368	8	1.1	299	2	O87489	O87489 pseudomonas
	369	8	1.1	304	10	O82321	O82321 arabidopsis
c	370	8	1.1	304	11	Q9EQN6	Q9eqn6 mus musculu
c	371	8	1.1	305	8	Q02677	Q02677 podospora a
c	372	8	1.1	307	5	Q9VAC9	Q9vac9 drosophila
c	373	8	1.1	320	10	O23160	O23160 arabidopsis
c	374	8	1.1	323	12	Q67712	Q67712 human adeno
	375	8	1.1	323	16	Q99UK0	Q99uk0 staphylococ
	376	8	1.1	323	16	Q8YTP3	Q8ytp3 anabaena sp
	377	8	1.1	324	10	Q9FQ84	Q9fq84 nicotiana t
c	378	8	1.1	324	10	Q942K9	Q942k9 oryza sativ
	379	8	1.1	327	10	Q94DI2	Q94di2 oryza sativ
	380	8	1.1	328	2	Q9L6R8	Q9l6r8 salmonella
c	381	8	1.1	335	2	Q8VUZ9	Q8vuz9 streptomyce
	382	8	1.1	337	17	Q96XQ7	Q96xq7 sulfolobus
c	383	8	1.1	341	16	Q9RTY3	Q9rty3 deinococcus
	384	8	1.1	345	16	Q92WS8	Q92ws8 rhizobium m
	385	8	1.1	345	16	Q8YB95	Q8yb95 brucella me
	386	8	1.1	346	16	Q8U6Y0	Q8u6y0 agrobacteri
c	387	8	1.1	346	16	Q9ACU5	Q9acu5 streptomyce
c	388	8	1.1	351	16	Q9PF30	Q9pf30 xylella fas
	389	8	1.1	353	4	Q8TEC9	Q8tec9 homo sapien
	390	8	1.1	360	10	Q9ZS03	Q9zs03 arabidopsis
c	391	8	1.1	362	10	Q9FJP8	Q9fjp8 arabidopsis
	392	8	1.1	365	4	Q9Y3V8	Q9y3v8 homo sapien
	393	8	1.1	365	10	Q9FQ86	Q9fq86 nicotiana t
c	394	8	1.1	368	16	Q92P72	Q92p72 rhizobium m
	395	8	1.1	369	4	Q9UHL0	Q9uhl0 homo sapien
	396	8	1.1	369	11	Q9QY16	Q9qy16 rattus norv
	397	8	1.1	369	11	Q9QY15	Q9qy15 mus musculu
	398	8	1.1	369	11	Q8R1B6	Q8r1b6 mus musculu
	399	8	1.1	370	4	Q96KE7	Q96ke7 homo sapien
c	400	8	1.1	372	11	Q9D7Z2	Q9d7z2 mus musculu
c	401	8	1.1	372	11	Q8R0U2	Q8r0u2 mus musculu
	402	8	1.1	372	13	Q90Y50	Q90y50 brachydanio
	403	8	1.1	373	17	Q978T9	Q978t9 thermoplasm
	404	8	1.1	374	16	Q97KY2	Q97ky2 clostridium
	405	8	1.1	375	5	Q9NGF5	Q9ngf5 drosophila
	406	8	1.1	375	5	Q9NGF4	Q9ngf4 drosophila
	407	8	1.1	375	5	Q9NGF3	Q9ngf3 drosophila
	408	8	1.1	375	5	Q9N6A0	Q9n6a0 drosophila
c	409	8	1.1	378	8	Q94ZK8	Q94zk8 cebus albif
c	410	8	1.1	379	5	O44879	O44879 caenorhabdi
	411	8	1.1	381	8	Q9SEV5	Q9sev5 guillardia
c	412	8	1.1	384	5	Q8SYB8	Q8syb8 drosophila
	413	8	1.1	388	17	Q9HLD1	Q9hld1 thermoplasm
c	414	8	1.1	389	10	Q94HS3	Q94hs3 oryza sativ
	415	8	1.1	390	10	Q9FQ90	Q9fq90 nicotiana t
c	416	8	1.1	394	17	Q8TQ10	Q8tq10 methanosarc
	417	8	1.1	398	11	Q9D8Q2	Q9d8q2 mus musculu
	418	8	1.1	399	4	Q9NTA9	Q9nta9 homo sapien
	419	8	1.1	400	5	Q24731	Q24731 drosophila
	420	8	1.1	407	16	Q9ZCQ0	Q9zcq0 rickettsia
	421	8	1.1	414	16	Q92GV2	Q92gv2 rickettsia
c	422	8	1.1	415	10	Q9FX26	Q9fx26 arabidopsis
c	423	8	1.1	419	13	Q8QG81	Q8qg81 xenopus lae
	424	8	1.1	420	4	Q9Y482	Q9y482 homo sapien
c	425	8	1.1	420	13	Q9DEG1	Q9deg1 gallus gall
c	426	8	1.1	423	10	Q9C531	Q9c531 arabidopsis
	427	8	1.1	425	10	Q9FQ91	Q9fq91 nicotiana t
	428	8	1.1	426	4	Q8WV76	Q8wv76 homo sapien
	429	8	1.1	428	16	Q9KKW0	Q9kkw0 vibrio chol
c	430	8	1.1	438	10	Q9CAU2	Q9cau2 arabidopsis
c	431	8	1.1	439	10	Q8VYM5	Q8vym5 arabidopsis
c	432	8	1.1	441	10	Q9CAU4	Q9cau4 arabidopsis
c	433	8	1.1	441	10	Q9CAU3	Q9cau3 arabidopsis
c	434	8	1.1	442	4	Q9HAI9	Q9hai9 homo sapien
	435	8	1.1	448	4	Q9H0U0	Q9h0u0 homo sapien
	436	8	1.1	450	10	O23251	O23251 arabidopsis
	437	8	1.1	451	10	Q9ZS07	Q9zs07 arabidopsis
c	438	8	1.1	453	16	P72834	P72834 synechocyst
c	439	8	1.1	457	10	Q9FF87	Q9ff87 arabidopsis
c	440	8	1.1	457	10	Q949W1	Q949w1 arabidopsis
c	441	8	1.1	461	10	Q9LTW4	Q9ltw4 arabidopsis
c	442	8	1.1	461	16	Q9CK50	Q9ck50 pasteurella
c	443	8	1.1	464	8	Q9B8C8	Q9b8c8 candida alb
	444	8	1.1	464	10	Q8VYS6	Q8vys6 arabidopsis
	445	8	1.1	465	5	Q8SRN8	Q8srn8 encephalito
	446	8	1.1	465	10	Q9FVV4	Q9fvv4 arabidopsis
	447	8	1.1	466	10	Q9FE48	Q9fe48 nicotiana t
	448	8	1.1	466	10	Q9FQ92	Q9fq92 nicotiana t
	449	8	1.1	467	4	Q9BWI8	Q9bwi8 homo sapien
	450	8	1.1	471	4	Q9HCH7	Q9hch7 homo sapien
	451	8	1.1	472	5	Q95PU7	Q95pu7 chironomus
	452	8	1.1	473	5	Q9Y0S9	Q9y0s9 drosophila
	453	8	1.1	474	10	O22679	O22679 arabidopsis
	454	8	1.1	476	16	Q9A9X6	Q9a9x6 caulobacter
	455	8	1.1	477	10	O65255	O65255 arabidopsis
	456	8	1.1	478	4	Q9HBZ9	Q9hbz9 homo sapien
	457	8	1.1	478	4	Q9NUU7	Q9nuu7 homo sapien
	458	8	1.1	478	11	Q921R0	Q921r0 mus musculu
	459	8	1.1	479	11	Q8R3C7	Q8r3c7 mus musculu
	460	8	1.1	480	16	Q98RE0	Q98re0 mycoplasma
	461	8	1.1	483	3	Q8X0X2	Q8x0x2 neurospora
c	462	8	1.1	483	5	Q8T0P0	Q8t0p0 drosophila
	463	8	1.1	483	13	Q9DGP9	Q9dgp9 xenopus lae
	464	8	1.1	487	2	Q9F0Z5	Q9f0z5 synechococc
c	465	8	1.1	488	5	Q9NBA7	Q9nba7 plasmodium
	466	8	1.1	491	4	Q9NVE0	Q9nve0 homo sapien
	467	8	1.1	491	16	Q9ZMI8	Q9zmi8 helicobacte
c	468	8	1.1	492	5	Q9NAW4	Q9naw4 plasmodium
	469	8	1.1	492	16	O25029	O25029 helicobacte
	470	8	1.1	492	16	Q55804	Q55804 synechocyst
	471	8	1.1	494	10	Q9SRW8	Q9srw8 arabidopsis
	472	8	1.1	496	10	Q9SZ89	Q9sz89 arabidopsis
c	473	8	1.1	496	16	Q9RRC0	Q9rrc0 deinococcus
	474	8	1.1	498	10	Q9ZS08	Q9zs08 arabidopsis
	475	8	1.1	498	10	Q9M2E0	Q9m2e0 arabidopsis
	476	8	1.1	499	10	O65275	O65275 arabidopsis
	477	8	1.1	499	11	Q9CS87	Q9cs87 mus musculu
c	478	8	1.1	502	12	Q8UYJ5	Q8uyj5 camelpox vi
	479	8	1.1	505	10	Q9ZS12	Q9zs12 arabidopsis
	480	8	1.1	505	10	Q8RXK6	Q8rxk6 arabidopsis
	481	8	1.1	506	16	Q99SH6	Q99sh6 staphylococ
	482	8	1.1	507	10	Q942N1	Q942n1 oryza sativ
	483	8	1.1	510	2	O07897	O07897 thermus the
c	484	8	1.1	513	13	Q9DEG2	Q9deg2 gallus gall
c	485	8	1.1	520	5	Q9BLY0	Q9bly0 ciona intes
	486	8	1.1	521	10	Q9FVY5	Q9fvy5 oryza sativ
	487	8	1.1	522	10	Q8VYG8	Q8vyg8 arabidopsis
c	488	8	1.1	524	5	Q9VFQ3	Q9vfq3 drosophila
c	489	8	1.1	526	4	Q8TBR3	Q8tbr3 homo sapien
	490	8	1.1	528	10	O80838	O80838 arabidopsis
	491	8	1.1	528	10	Q94BV4	Q94bv4 arabidopsis
	492	8	1.1	528	11	Q9CSI0	Q9csi0 mus musculu
c	493	8	1.1	528	11	Q8R4H7	Q8r4h7 mus musculu
	494	8	1.1	535	3	O13622	O13622 schizosacch
c	495	8	1.1	543	4	Q9P1Y9	Q9p1y9 homo sapien
	496	8	1.1	544	5	O61815	O61815 caenorhabdi
	497	8	1.1	544	16	Q9PA24	Q9pa24 xylella fas
	498	8	1.1	546	5	O61097	O61097 trypanosoma
	499	8	1.1	546	5	O15919	O15919 trypanosoma
c	500	8	1.1	550	13	O93263	O93263 gallus gall
c	501	8	1.1	552	5	Q9VNS2	Q9vns2 drosophila
c	502	8	1.1	552	10	Q94LG7	Q94lg7 oryza sativ
c	503	8	1.1	557	3	O93981	O93981 gibberella
	504	8	1.1	558	10	Q9C9H4	Q9c9h4 arabidopsis
c	505	8	1.1	561	11	Q9CXH6	Q9cxh6 mus musculu
	506	8	1.1	566	5	Q23910	Q23910 dictyosteli
	507	8	1.1	568	10	Q9LIH9	Q9lih9 arabidopsis
	508	8	1.1	568	11	Q9D2M2	Q9d2m2 mus musculu
	509	8	1.1	574	3	O93990	O93990 candida alb
c	510	8	1.1	588	10	Q9LT69	Q9lt69 arabidopsis
	511	8	1.1	591	10	Q942N2	Q942n2 oryza sativ
	512	8	1.1	593	10	Q9FLB0	Q9flb0 arabidopsis

	513	8	1.1	594	5	Q9VVK8	Q9vvk8 drosophila
	514	8	1.1	595	3	O60080	O60080 schizosacch
	515	8	1.1	598	4	Q9Y2R4	Q9y2r4 homo sapien
	516	8	1.1	598	11	Q99PT0	Q99pt0 rattus norv
	517	8	1.1	600	3	O93935	O93935 paracoccidi
c	518	8	1.1	603	5	Q9VW86	Q9vw86 drosophila
	519	8	1.1	604	3	O74491	O74491 schizosacch
	520	8	1.1	606	3	O74764	O74764 schizosacch
	521	8	1.1	613	5	O62372	O62372 caenorhabdi
	522	8	1.1	613	5	Q9VHU1	Q9vhu1 drosophila
	523	8	1.1	615	16	Q8U8C1	Q8u8c1 agrobacteri
c	524	8	1.1	618	13	Q90W76	Q90w76 oncorhynchu
	525	8	1.1	620	16	Q98E62	Q98e62 rhizobium l
	526	8	1.1	621	10	Q94IV5	Q94iv5 oryza sativ
	527	8	1.1	626	10	Q9SW44	Q9sw44 arabidopsis
	528	8	1.1	626	16	Q8YEQ4	Q8yeq4 brucella me
	529	8	1.1	630	16	Q92SG6	Q92sg6 rhizobium m
	530	8	1.1	638	5	O18056	O18056 caenorhabdi
	531	8	1.1	641	5	Q95SF2	Q95sf2 drosophila
c	532	8	1.1	643	5	Q8ST22	Q8st22 dictyosteli
	533	8	1.1	647	10	Q942Q0	Q942q0 oryza sativ
	534	8	1.1	648	3	O74393	O74393 schizosacch
c	535	8	1.1	659	5	Q94600	Q94600 leishmania
	536	8	1.1	660	11	Q9D2E0	Q9d2e0 mus musculu
	537	8	1.1	663	16	Q9KLE2	Q9kle2 vibrio chol
c	538	8	1.1	665	11	Q9QXG2	Q9qxg2 mus musculu
	539	8	1.1	669	11	Q9JIH4	Q9jih4 mus musculu
	540	8	1.1	671	10	Q39189	Q39189 arabidopsis
	541	8	1.1	671	10	Q9LVB7	Q9lvb7 arabidopsis
	542	8	1.1	671	10	Q94BX7	Q94bx7 arabidopsis
	543	8	1.1	680	5	Q8SYP5	Q8syp5 drosophila
c	544	8	1.1	680	10	Q9SY66	Q9sy66 arabidopsis
	545	8	1.1	683	10	O23506	O23506 arabidopsis
c	546	8	1.1	684	4	Q8TC67	Q8tc67 homo sapien
	547	8	1.1	685	10	Q41382	Q41382 spinacia ol
c	548	8	1.1	690	4	Q8WVZ6	Q8wvz6 homo sapien
	549	8	1.1	691	10	Q9FK52	Q9fk52 arabidopsis
	550	8	1.1	698	10	Q8S1J0	Q8s1j0 oryza sativ
	551	8	1.1	699	10	Q8W1K7	Q8w1k7 chlamydomon
	552	8	1.1	701	10	Q9SQR8	Q9sqr8 arabidopsis
	553	8	1.1	703	10	Q9LID3	Q9lid3 arabidopsis
	554	8	1.1	709	3	O60173	O60173 schizosacch
	555	8	1.1	713	10	Q9M6R6	Q9m6r6 phaseolus a
	556	8	1.1	715	3	Q8TFL3	Q8tfl3 candida gla
	557	8	1.1	727	5	Q9VNV3	Q9vnv3 drosophila
	558	8	1.1	734	11	Q99MJ9	Q99mj9 mus musculu
	559	8	1.1	737	4	Q9BQ39	Q9bq39 homo sapien
	560	8	1.1	748	10	Q9SB96	Q9sb96 arabidopsis
	561	8	1.1	748	10	Q8RUW1	Q8ruw1 arabidopsis
c	562	8	1.1	755	10	Q9FP21	Q9fp21 oryza sativ
	563	8	1.1	759	13	Q9DF35	Q9df35 xenopus lae
c	564	8	1.1	766	10	Q8W0Z5	Q8w0z5 arabidopsis
c	565	8	1.1	784	13	Q90YB5	Q90yb5 gallus gall
c	566	8	1.1	787	10	O82362	O82362 arabidopsis
	567	8	1.1	797	10	Q93Y39	Q93y39 arabidopsis
	568	8	1.1	800	13	Q9DF36	Q9df36 xenopus lae
	569	8	1.1	802	5	Q9VJ74	Q9vj74 drosophila
	570	8	1.1	805	5	Q9VX34	Q9vx34 drosophila
	571	8	1.1	805	11	Q9CWZ5	Q9cwz5 mus musculu
c	572	8	1.1	807	10	Q8RWX6	Q8rwx6 arabidopsis
c	573	8	1.1	810	3	Q9P6V7	Q9p6v7 neurospora
	574	8	1.1	813	5	Q9V915	Q9v915 drosophila
	575	8	1.1	821	17	Q8TUH1	Q8tuh1 methanosarc
	576	8	1.1	832	10	Q9LRY9	Q9lry9 arabidopsis
c	577	8	1.1	850	11	Q9QWH1	Q9qwh1 mus musculu
	578	8	1.1	853	16	Q930Y7	Q930y7 rhizobium m
	579	8	1.1	879	16	Q9FBJ2	Q9fbj2 streptomyce
c	580	8	1.1	883	17	Q8TTT9	Q8ttt9 methanosarc
	581	8	1.1	920	3	O74429	O74429 schizosacch
c	582	8	1.1	926	4	O95462	O95462 homo sapien
c	583	8	1.1	926	4	Q9UDR5	Q9udr5 homo sapien
	584	8	1.1	934	10	Q94HQ0	Q94hq0 oryza sativ
c	585	8	1.1	945	5	Q9NKT7	Q9nkt7 leishmania

c	586	8	1.1	948	5	Q8T065	Q8t065 drosophila
	587	8	1.1	953	4	Q14151	Q14151 homo sapien
c	588	8	1.1	965	10	Q9MA56	Q9ma56 arabidopsis
	589	8	1.1	1005	5	Q95ZC5	Q95zc5 leishmania
	590	8	1.1	1022	5	Q22308	Q22308 caenorhabdi
c	591	8	1.1	1096	5	Q9W117	Q9w117 drosophila
c	592	8	1.1	1108	5	Q9V788	Q9v788 drosophila
c	593	8	1.1	1164	5	Q9VI61	Q9vi61 drosophila
c	594	8	1.1	1272	5	Q95SG4	Q95sg4 drosophila
	595	8	1.1	1283	5	Q8T5H0	Q8t5h0 anopheles g
c	596	8	1.1	1283	5	Q8T5H0	Q8t5h0 anopheles g
c	597	8	1.1	1338	5	Q9V416	Q9v416 drosophila
c	598	8	1.1	1368	5	Q9VHH9	Q9vhh9 drosophila
c	599	8	1.1	1432	10	Q9FPR8	Q9fpr8 chlamydomon
c	600	8	1.1	1443	5	Q9N8S1	Q9n8s1 trypanosoma
c	601	8	1.1	1458	5	Q8T9F7	Q8t9f7 drosophila
	602	8	1.1	1458	11	Q91WY0	Q91wy0 mus musculu
	603	8	1.1	1458	11	Q8VHJ6	Q8vhj6 mus musculu
c	604	8	1.1	1478	17	Q8TLD5	Q8tld5 methanosarc
	605	8	1.1	1484	4	Q96QH3	Q96qh3 homo sapien
	606	8	1.1	1487	11	Q9JLA1	Q9jla1 mus musculu
c	607	8	1.1	1661	6	Q9GKP1	Q9gkp1 sus scrofa
	608	8	1.1	1665	16	O53215	O53215 mycobacteri
c	609	8	1.1	1783	5	Q9VX48	Q9vx48 drosophila
c	610	8	1.1	1847	5	Q9NKN5	Q9nkn5 leishmania
c	611	8	1.1	1957	11	Q9JMJ9	Q9jmj9 mus musculu
c	612	8	1.1	2083	11	Q60997	Q60997 mus musculu
c	613	8	1.1	2090	5	Q9W2T1	Q9w2t1 drosophila
c	614	8	1.1	2157	4	O95875	O95875 homo sapien
c	615	8	1.1	2157	4	Q96QC6	Q96qc6 homo sapien
c	616	8	1.1	2487	6	Q9N1T0	Q9n1t0 ornithorhyn
	617	8	1.1	3034	12	Q8V1Y4	Q8v1y4 wheat strea
	618	8	1.1	3035	12	O72347	O72347 wheat strea
	619	8	1.1	3035	12	Q9E161	Q9e161 wheat strea
	620	8	1.1	3035	12	Q8V1Y3	Q8v1y3 wheat strea
c	621	8	1.1	3389	4	Q96QU9	Q96qu9 homo sapien
c	622	8	1.1	3508	4	Q96RM4	Q96rm4 homo sapien
c	623	8	1.1	3616	13	Q9W6V0	Q9w6v0 gallus gall
c	624	7	1.0	20	11	Q9QUU8	Q9quu8 mus sp. c-m
	625	7	1.0	29	2	Q9RCT6	Q9rct6 streptococc
	626	7	1.0	31	2	Q9RCT5	Q9rct5 streptococc
c	627	7	1.0	31	15	O11551	O11551 human immun
	628	7	1.0	33	2	P71125	P71125 campylobact
	629	7	1.0	35	2	Q9RCT4	Q9rct4 streptococc
c	630	7	1.0	35	4	Q13264	Q13264 homo sapien
c	631	7	1.0	37	4	Q8TD81	Q8td81 homo sapien
	632	7	1.0	39	2	Q9RCT8	Q9rct8 streptococc
c	633	7	1.0	42	16	Q8X9U0	Q8x9u0 escherichia
	634	7	1.0	45	16	O50794	O50794 borrelia bu
	635	7	1.0	51	16	Q92K67	Q92k67 rhizobium m
c	636	7	1.0	54	10	Q9S8X3	Q9s8x3 triticum ae
c	637	7	1.0	54	10	Q9S8X2	Q9s8x2 triticum ae
	638	7	1.0	55	2	Q9RCT7	Q9rct7 streptococc
	639	7	1.0	58	2	Q9RCT9	Q9rct9 streptococc
c	640	7	1.0	63	2	Q9AN47	Q9an47 bradyrhizob
c	641	7	1.0	63	8	O20444	O20444 limulus pol
	642	7	1.0	63	16	Q8X5E3	Q8x5e3 escherichia
	643	7	1.0	65	5	Q9GYI9	Q9gyi9 caenorhabdi
c	644	7	1.0	65	16	Q99TK2	Q99tk2 staphylococ
c	645	7	1.0	66	16	Q98PB9	Q98pb9 rhizobium l
c	646	7	1.0	68	2	Q52624	Q52624 proteus vul
	647	7	1.0	68	10	Q9S967	Q9s967 medicago sa
c	648	7	1.0	69	10	Q94K36	Q94k36 arabidopsis
c	649	7	1.0	70	12	Q8UXA5	Q8uxa5 hepatitis c
c	650	7	1.0	70	12	Q8UXA3	Q8uxa3 hepatitis c
c	651	7	1.0	70	12	Q8UX97	Q8ux97 hepatitis c
c	652	7	1.0	70	12	Q8UX94	Q8ux94 hepatitis c
c	653	7	1.0	70	12	Q8UX89	Q8ux89 hepatitis c
c	654	7	1.0	70	12	Q8UX86	Q8ux86 hepatitis c
c	655	7	1.0	70	12	Q8UX84	Q8ux84 hepatitis c
c	656	7	1.0	70	12	Q8UX83	Q8ux83 hepatitis c
c	657	7	1.0	70	12	Q8UX82	Q8ux82 hepatitis c
c	658	7	1.0	70	12	Q8UX80	Q8ux80 hepatitis c

c	659	7	1.0	70	12	Q8UX78	Q8ux78 hepatitis c
c	660	7	1.0	70	12	Q8UX76	Q8ux76 hepatitis c
c	661	7	1.0	70	12	Q8UX74	Q8ux74 hepatitis c
c	662	7	1.0	70	12	Q8UX69	Q8ux69 hepatitis c
c	663	7	1.0	70	12	Q8UX67	Q8ux67 hepatitis c
c	664	7	1.0	70	12	Q8UX66	Q8ux66 hepatitis c
c	665	7	1.0	70	12	Q8UX65	Q8ux65 hepatitis c
	666	7	1.0	71	4	Q9NPY6	Q9npy6 homo sapien
c	667	7	1.0	71	17	Q8TSL0	Q8tsl0 methanosarc
	668	7	1.0	71	17	Q8TPF3	Q8tpf3 methanosarc
c	669	7	1.0	72	6	Q28795	Q28795 tayassu taj
c	670	7	1.0	72	16	Q9KY33	Q9ky33 streptomyce
c	671	7	1.0	73	5	Q9VIQ3	Q9viq3 drosophila
c	672	7	1.0	75	5	Q9NLM8	Q9nlm8 leishmania
c	673	7	1.0	78	2	Q51924	Q51924 pasteurella
	674	7	1.0	80	10	O04126	O04126 zea mays (m
c	675	7	1.0	83	16	Q8UEM9	Q8uem9 agrobacteri
	676	7	1.0	85	16	Q9A0I5	Q9a0i5 streptococc
	677	7	1.0	85	16	Q98ML7	Q98ml7 rhizobium l
	678	7	1.0	85	16	Q9RD93	Q9rd93 streptomyce
	679	7	1.0	86	8	Q9B8X2	Q9b8x2 schistosoma
	680	7	1.0	86	12	Q919M4	Q919m4 culex nigri
	681	7	1.0	90	16	O83111	O83111 treponema p
c	682	7	1.0	91	13	Q9YHL0	Q9yhl0 ginglymosto
c	683	7	1.0	91	16	Q9JX58	Q9jx58 neisseria m
	684	7	1.0	91	17	Q8ZWQ1	Q8zwq1 pyrobaculum
	685	7	1.0	92	16	Q9ZKI7	Q9zki7 helicobacte
c	686	7	1.0	93	10	Q9AUU6	Q9auu6 oryza sativ
	687	7	1.0	95	5	Q00088	Q00088 drosophila
c	688	7	1.0	96	2	Q9Z9I4	Q9z9i4 streptomyce
c	689	7	1.0	96	5	Q9VKS2	Q9vks2 drosophila
c	690	7	1.0	96	10	Q9SDK2	Q9sdk2 oryza sativ
c	691	7	1.0	96	10	Q94HN5	Q94hn5 oryza sativ
	692	7	1.0	98	5	Q23908	Q23908 dictyosteli
c	693	7	1.0	98	5	Q9VCP2	Q9vcp2 drosophila
	694	7	1.0	98	12	Q8QXX9	Q8qxx9 garlic late
c	695	7	1.0	99	16	Q8YRF9	Q8yrf9 anabaena sp
	696	7	1.0	100	8	Q952Q7	Q952q7 lycopersico
	697	7	1.0	100	8	Q952Q6	Q952q6 podophyllum
c	698	7	1.0	104	10	Q9LRN7	Q9lrn7 arabidopsis
c	699	7	1.0	104	16	O31480	O31480 bacillus su
c	700	7	1.0	104	16	Q8XV14	Q8xv14 ralstonia s
c	701	7	1.0	105	11	Q99J74	Q99j74 mus musculu
	702	7	1.0	105	16	Q99S32	Q99s32 staphylococ
	703	7	1.0	106	2	O05681	O05681 mycobacteri
c	704	7	1.0	106	10	Q9FUU0	Q9fuu0 oxalis regn
	705	7	1.0	107	5	O97028	O97028 dugesia jap
c	706	7	1.0	107	17	Q96Z20	Q96z20 sulfolobus
	707	7	1.0	108	5	O97023	O97023 dugesia jap
c	708	7	1.0	108	8	Q94V73	Q94v73 aricia ages
c	709	7	1.0	108	8	Q94PT7	Q94pt7 aricia ages
c	710	7	1.0	109	5	Q9W486	Q9w486 drosophila
	711	7	1.0	109	16	Q9A375	Q9a375 caulobacter
c	712	7	1.0	109	16	Q932B9	Q932b9 staphylococ
c	713	7	1.0	111	2	O85208	O85208 vibrio chol
	714	7	1.0	111	5	O97024	O97024 dugesia jap
c	715	7	1.0	111	16	Q8Y0B4	Q8y0b4 ralstonia s
c	716	7	1.0	112	3	Q07255	Q07255 saccharomyc
c	717	7	1.0	112	17	Q97ZU5	Q97zu5 sulfolobus
	718	7	1.0	113	5	O15720	O15720 dictyosteli
	719	7	1.0	113	17	Q96XQ9	Q96xq9 sulfolobus
c	720	7	1.0	114	10	Q9ST83	Q9st83 oryza sativ
c	721	7	1.0	116	17	Q9YDL6	Q9ydl6 aeropyrum p
	722	7	1.0	118	5	Q21028	Q21028 caenorhabdi
c	723	7	1.0	120	5	O45262	O45262 caenorhabdi
c	724	7	1.0	121	16	Q9PFV3	Q9pfv3 xylella fas
c	725	7	1.0	122	5	O44205	O44205 megaselia s
c	726	7	1.0	124	10	O49587	O49587 arabidopsis
c	727	7	1.0	124	12	Q9DKM1	Q9dkm1 spodoptera
	728	7	1.0	125	3	Q96VH3	Q96vh3 saccharomyc
c	729	7	1.0	126	10	O64723	O64723 arabidopsis
c	730	7	1.0	126	16	Q97TQ7	Q97tq7 clostridium
c	731	7	1.0	127	17	Q9Y9E2	Q9y9e2 aeropyrum p

c	732	7	1.0	128	12	Q68173	Q68173 hepatitis c
c	733	7	1.0	129	8	Q9B130	Q9b130 aedes aegyp
c	734	7	1.0	129	8	Q9B0U4	Q9b0u4 aedes aegyp
c	735	7	1.0	129	8	Q9B0U3	Q9b0u3 aedes aegyp
c	736	7	1.0	129	8	Q9B0U2	Q9b0u2 aedes aegyp
c	737	7	1.0	129	8	Q9B3H4	Q9b3h4 aedes aegyp
c	738	7	1.0	129	8	Q9B3H3	Q9b3h3 aedes aegyp
c	739	7	1.0	129	8	Q9B3H2	Q9b3h2 aedes aegyp
c	740	7	1.0	129	8	Q9B3H1	Q9b3h1 aedes aegyp
c	741	7	1.0	129	8	Q9B3H0	Q9b3h0 aedes aegyp
c	742	7	1.0	129	8	Q9B3G9	Q9b3g9 aedes aegyp
c	743	7	1.0	129	8	Q9B3G8	Q9b3g8 aedes aegyp
c	744	7	1.0	129	8	Q9B3G7	Q9b3g7 aedes aegyp
c	745	7	1.0	129	10	Q9SXW1	Q9sxw1 lithospermu
c	746	7	1.0	129	12	Q8VBA0	Q8vba0 white spot
c	747	7	1.0	130	8	Q9MDS0	Q9mds0 aedes aegyp
c	748	7	1.0	130	8	Q9MDR9	Q9mdr9 aedes aegyp
	749	7	1.0	130	16	Q9ZC14	Q9zc14 streptomyce
	750	7	1.0	131	11	Q99LZ7	Q99lz7 mus musculu
c	751	7	1.0	131	15	Q88389	Q88389 primate t-l
c	752	7	1.0	131	17	Q8TW52	Q8tw52 methanopyru
c	753	7	1.0	132	6	Q9GMU5	Q9gmu5 macaca fasc
	754	7	1.0	132	16	Q8XD66	Q8xd66 escherichia
	755	7	1.0	133	11	Q921B5	Q921b5 mus musculu
c	756	7	1.0	134	10	Q40447	Q40447 nicotiana g
	757	7	1.0	134	10	Q8S4D4	Q8s4d4 podococcus
c	758	7	1.0	135	2	P70762	P70762 anabaena sp
c	759	7	1.0	135	5	O44206	O44206 megaselia s
c	760	7	1.0	135	5	Q9VUG0	Q9vug0 drosophila
c	761	7	1.0	135	6	Q9GMC4	Q9gmc4 ovis aries
c	762	7	1.0	135	15	Q80829	Q80829 human t-cel
	763	7	1.0	136	17	Q9HMG2	Q9hmg2 halobacteri
c	764	7	1.0	138	11	Q9D3Y0	Q9d3y0 mus musculu
c	765	7	1.0	138	15	Q80823	Q80823 human t-cel
c	766	7	1.0	139	16	Q8U5I9	Q8u5i9 agrobacteri
c	767	7	1.0	140	16	Q99RP2	Q99rp2 staphylococ
	768	7	1.0	141	5	Q967N1	Q967n1 micracis ca
	769	7	1.0	141	6	Q9BER7	Q9ber7 leopardus p
c	770	7	1.0	142	2	P82552	P82552 streptococc
c	771	7	1.0	142	4	Q8WW60	Q8ww60 homo sapien
c	772	7	1.0	143	3	O74666	O74666 rhizopus or
c	773	7	1.0	143	5	Q18333	Q18333 caenorhabdi
	774	7	1.0	143	6	Q9BER6	Q9ber6 panthera on
c	775	7	1.0	144	5	O16707	O16707 caenorhabdi
c	776	7	1.0	144	5	Q9GZ13	Q9gz13 caenorhabdi
c	777	7	1.0	144	10	Q9FME9	Q9fme9 arabidopsis
c	778	7	1.0	144	16	Q9PC82	Q9pc82 xylella fas
	779	7	1.0	144	16	Q9A7T0	Q9a7t0 caulobacter
	780	7	1.0	145	10	P93293	P93293 arabidopsis
c	781	7	1.0	146	5	Q9XZT7	Q9xzt7 drosophila
c	782	7	1.0	146	8	Q9TAN9	Q9tan9 australophi
	783	7	1.0	146	16	Q97NF2	Q97nf2 streptococc
c	784	7	1.0	147	10	Q9M7Y1	Q9m7y1 arabidopsis
c	785	7	1.0	147	11	Q9D1V3	Q9d1v3 mus musculu
	786	7	1.0	148	5	Q962A0	Q962a0 hylastes po
	787	7	1.0	148	5	Q8SWM0	Q8swm0 encephalito
	788	7	1.0	148	6	Q9BEV1	Q9bev1 talpa altai
	789	7	1.0	148	6	Q9BEV0	Q9bev0 condylura c
	790	7	1.0	148	6	Q9BEU9	Q9beu9 sorex arane
	791	7	1.0	148	6	Q9BEU3	Q9beu3 orycteropus
	792	7	1.0	148	6	Q9BER9	Q9ber9 tapirus ind
	793	7	1.0	148	6	Q9BER8	Q9ber8 felis silve
	794	7	1.0	148	6	Q9BER5	Q9ber5 ursus arcto
	795	7	1.0	148	11	Q99NM3	Q99nm3 pedetes cap
	796	7	1.0	148	11	Q99NL7	Q99nl7 dipodomys h
	797	7	1.0	148	11	Q99NL5	Q99nl5 cavia tschu
	798	7	1.0	148	11	Q99NL4	Q99nl4 hydrochoeru
	799	7	1.0	148	11	Q99NL2	Q99nl2 dinomys bra
	800	7	1.0	149	5	Q967M8	Q967m8 premnobius
	801	7	1.0	149	5	Q963I1	Q963i1 dendroctonu
	802	7	1.0	149	5	Q961Z4	Q961z4 dendroctonu
	803	7	1.0	149	5	Q961Z0	Q961z0 hylurgus mi
c	804	7	1.0	149	11	Q9D7P2	Q9d7p2 mus musculu

c 805	7	1.0	149	12	Q68719	Q68719 hepatitis c
c 806	7	1.0	149	15	Q82393	Q82393 human t-cel
c 807	7	1.0	149	15	Q82395	Q82395 human t-cel
c 808	7	1.0	149	15	Q82397	Q82397 human t-cel
c 809	7	1.0	149	15	Q82399	Q82399 human t-cel
c 810	7	1.0	149	15	Q82405	Q82405 human t-cel
c 811	7	1.0	149	15	Q82408	Q82408 human t-cel
812	7	1.0	150	5	Q961Z2	Q961z2 hylurdrecto
c 813	7	1.0	150	10	O80490	O80490 arabidopsis
c 814	7	1.0	150	13	Q90ZI4	Q90zi4 pagrus majo
815	7	1.0	150	13	Q9PSH8	Q9psh8 xenopus lae
816	7	1.0	150	16	Q9L0P6	Q9l0p6 streptomyce
817	7	1.0	151	11	Q99PU1	Q99pu1 mus musculu
c 818	7	1.0	151	12	O12669	O12669 colobus mon
c 819	7	1.0	151	15	Q97263	Q97263 human immun
820	7	1.0	152	5	Q961Z8	Q961z8 dendroctonu
821	7	1.0	153	5	Q961Z3	Q961z3 hylurdrecto
822	7	1.0	153	16	Q8YSA6	Q8ysa6 anabaena sp
c 823	7	1.0	153	17	Q9YFD4	Q9yfd4 aeropyrum p
c 824	7	1.0	154	5	Q9VET5	Q9vet5 drosophila
825	7	1.0	155	5	Q961Z1	Q961z1 hylurgonotu
c 826	7	1.0	155	5	Q9VZK6	Q9vzk6 drosophila
c 827	7	1.0	155	8	Q9G7M4	Q9g7m4 ochlerotatu
828	7	1.0	155	11	Q99PT6	Q99pt6 mus musculu
829	7	1.0	155	16	Q97G78	Q97g78 clostridium
830	7	1.0	156	5	Q961Z7	Q961z7 dendroctonu
831	7	1.0	157	10	Q9AVD9	Q9avd9 nicotiana t
832	7	1.0	157	10	Q9AVD2	Q9avd2 nicotiana t
833	7	1.0	157	10	Q9AVD1	Q9avd1 nicotiana t
834	7	1.0	157	10	Q40238	Q40238 lycopersico
835	7	1.0	158	5	Q963H5	Q963h5 pityokteine
836	7	1.0	158	5	Q963H4	Q963h4 pityophthor
837	7	1.0	158	10	Q9AVD8	Q9avd8 nicotiana t
838	7	1.0	158	10	Q9FQL8	Q9fql8 cyanophora
839	7	1.0	158	10	Q9ZW37	Q9zw37 arabidopsis
c 840	7	1.0	159	3	Q12135	Q12135 saccharomyc
c 841	7	1.0	159	10	O65683	O65683 arabidopsis
c 842	7	1.0	159	16	Q9RVK2	Q9rvk2 deinococcus
c 843	7	1.0	160	10	P94060	P94060 arabidopsis
844	7	1.0	160	11	O35058	O35058 mus musculu
845	7	1.0	161	5	Q963H7	Q963h7 hypothenemu
c 846	7	1.0	161	16	Q8ZB19	Q8zb19 yersinia pe
c 847	7	1.0	161	16	Q8Y9R1	Q8y9r1 listeria mo
c 848	7	1.0	161	17	Q8TLN3	Q8tln3 methanosarc
849	7	1.0	162	5	O61829	O61829 caenorhabdi
850	7	1.0	163	2	Q8VLS9	Q8vls9 wolbachia s
851	7	1.0	163	11	O54719	O54719 rattus norv
c 852	7	1.0	163	12	Q8VB89	Q8vb89 white spot
c 853	7	1.0	164	6	P79361	P79361 ovis aries
c 854	7	1.0	164	10	O48567	O48567 euphorbia e
c 855	7	1.0	165	11	Q9CPU3	Q9cpu3 mus musculu
856	7	1.0	165	11	O35067	O35067 mus musculu
857	7	1.0	165	16	Q8XB58	Q8xb58 escherichia
858	7	1.0	166	11	O35065	O35065 mus musculu
859	7	1.0	167	10	Q9AVD4	Q9avd4 nicotiana t
860	7	1.0	167	10	Q9AVD3	Q9avd3 nicotiana t
861	7	1.0	167	10	O64835	O64835 arabidopsis
c 862	7	1.0	167	16	Q9PCM7	Q9pcm7 xylella fas
c 863	7	1.0	168	5	P91969	P91969 trichinella
864	7	1.0	168	11	O35068	O35068 mus musculu
c 865	7	1.0	168	15	Q9DKR6	Q9dkr6 human t-cel
c 866	7	1.0	168	16	Q9RV18	Q9rv18 deinococcus
867	7	1.0	168	16	Q98GN6	Q98gn6 rhizobium l
868	7	1.0	169	5	Q9XUJ8	Q9xuj8 caenorhabdi
869	7	1.0	169	11	O35061	O35061 mus musculu
870	7	1.0	170	10	Q9AVD7	Q9avd7 nicotiana t
c 871	7	1.0	170	15	O70642	O70642 simian t-ly
c 872	7	1.0	170	15	Q9WI16	Q9wi16 human t-cel
c 873	7	1.0	170	15	Q9DKS0	Q9dks0 human t-cel
c 874	7	1.0	170	15	O73454	O73454 human t-cel
c 875	7	1.0	170	15	Q85601	Q85601 human t-cel
c 876	7	1.0	170	15	Q08838	Q08838 human t-cel
c 877	7	1.0	171	12	Q83420	Q83420 pseudorabie

c 878	7	1.0	174	6	Q8WNT5	Q8wnt5 canis lupus
c 879	7	1.0	174	16	Q9I4R7	Q9i4r7 pseudomonas
880	7	1.0	174	16	O86519	O86519 streptomyce
881	7	1.0	174	16	Q9R7V6	Q9r7v6 streptomyce
882	7	1.0	175	10	O65088	O65088 picea maria
c 883	7	1.0	175	11	Q9CYS3	Q9cys3 mus musculu
c 884	7	1.0	175	16	Q8XFS7	Q8xfs7 salmonella
885	7	1.0	176	2	Q9RCU0	Q9rcu0 streptococc
886	7	1.0	176	14	Q99IT7	Q99it7 uncultured
c 887	7	1.0	177	2	Q9KIH6	Q9kih6 rhizobium e
888	7	1.0	177	10	O48838	O48838 arabidopsis
c 889	7	1.0	177	10	Q9AV52	Q9av52 oryza sativ
890	7	1.0	177	10	Q8RWG6	Q8rwg6 arabidopsis
891	7	1.0	177	16	Q8XA52	Q8xa52 escherichia
c 892	7	1.0	178	5	Q8SXD7	Q8sxd7 drosophila
c 893	7	1.0	178	11	Q91X12	Q91x12 cavia porce
894	7	1.0	178	16	Q9PBT1	Q9pbt1 xylella fas
895	7	1.0	180	5	O97222	O97222 plasmodium
c 896	7	1.0	180	5	Q9VC28	Q9vc28 drosophila
897	7	1.0	180	10	Q9M9M0	Q9m9m0 arabidopsis
c 898	7	1.0	180	17	Q8U494	Q8u494 pyrococcus
c 899	7	1.0	181	5	Q9GZ14	Q9gz14 caenorhabdi
900	7	1.0	181	5	Q9W3J3	Q9w3j3 drosophila
c 901	7	1.0	181	5	Q9VWM4	Q9vwm4 drosophila
902	7	1.0	181	10	Q9SB89	Q9sb89 arabidopsis
903	7	1.0	181	16	Q9X833	Q9x833 streptomyce
c 904	7	1.0	182	10	Q948F5	Q948f5 oryza sativ
905	7	1.0	182	16	Q9KZR9	Q9kzr9 streptomyce
906	7	1.0	183	5	Q961Y9	Q961y9 pachycotes
907	7	1.0	183	16	Q9L140	Q9l140 streptomyce
908	7	1.0	184	12	Q89571	Q89571 african swi
909	7	1.0	184	16	Q9L235	Q9l235 streptomyce
910	7	1.0	186	2	Q56745	Q56745 wolbachia s
c 911	7	1.0	186	10	Q9FIF4	Q9fif4 arabidopsis
c 912	7	1.0	186	10	Q94EF9	Q94ef9 oryza sativ
913	7	1.0	186	10	Q8SA76	Q8sa76 oryza sativ
914	7	1.0	187	17	O27994	O27994 archaeoglob
915	7	1.0	188	10	O82214	O82214 arabidopsis
c 916	7	1.0	188	10	Q9SJK7	Q9sjk7 arabidopsis
c 917	7	1.0	188	11	Q9D8I1	Q9d8i1 mus musculu
c 918	7	1.0	189	10	Q94J56	Q94j56 oryza sativ
919	7	1.0	190	5	Q962A1	Q962a1 hylurgops r
c 920	7	1.0	190	6	Q9XSQ9	Q9xsq9 capra hircu
921	7	1.0	191	4	Q9H3E4	Q9h3e4 homo sapien
c 922	7	1.0	191	5	O46029	O46029 chironomus
923	7	1.0	191	5	Q9W2Z2	Q9w2z2 drosophila
924	7	1.0	191	16	Q9XAE1	Q9xae1 streptomyce
c 925	7	1.0	191	17	Q8ZUC8	Q8zuc8 pyrobaculum
926	7	1.0	191	17	Q8TZB0	Q8tzb0 methanopyru
927	7	1.0	192	10	Q9FIT9	Q9fit9 arabidopsis
c 928	7	1.0	192	10	Q9FW23	Q9fw23 oryza sativ
c 929	7	1.0	192	12	Q81414	Q81414 hepatitis c
c 930	7	1.0	192	12	Q81420	Q81420 hepatitis c
931	7	1.0	193	3	O42623	O42623 magnaporthe
c 932	7	1.0	193	5	Q8SY22	Q8sy22 drosophila
c 933	7	1.0	195	5	Q8SZG9	Q8szg9 drosophila
934	7	1.0	195	15	Q9DV50	Q9dv50 human immun
c 935	7	1.0	196	2	Q49632	Q49632 mycobacteri
936	7	1.0	197	2	O68889	O68889 lactococcus
937	7	1.0	198	5	Q9VPW6	Q9vpw6 drosophila
938	7	1.0	199	4	Q96EY2	Q96ey2 homo sapien
c 939	7	1.0	199	5	Q8T2K7	Q8t2k7 dictyosteli
940	7	1.0	199	11	Q922C9	Q922c9 mus musculu
c 941	7	1.0	200	5	Q9VEN6	Q9ven6 drosophila
c 942	7	1.0	201	5	Q9VJ78	Q9vj78 drosophila
943	7	1.0	201	16	Q9KZN3	Q9kzn3 streptomyce
944	7	1.0	202	10	Q9FEW8	Q9few8 lycopersico
945	7	1.0	202	10	Q93W87	Q93w87 arabidopsis
c 946	7	1.0	202	16	Q8XDI6	Q8xdi6 escherichia
947	7	1.0	202	17	Q8U1P5	Q8u1p5 pyrococcus
c 948	7	1.0	202	17	Q8TZ73	Q8tz73 methanopyru
949	7	1.0	204	2	Q9AJP1	Q9ajp1 streptomyce
950	7	1.0	204	5	Q963H6	Q963h6 pityogenes

```
   951    7   1.0   204   5  Q9VPI4          Q9vpi4 drosophila
   952    7   1.0   204  10  Q9LGY5          Q9lgy5 oryza sativ
   953    7   1.0   204  16  Q8XBM7          Q8xbm7 escherichia
   954    7   1.0   204  16  Q9KZI8          Q9kzi8 streptomyce
   955    7   1.0   205  17  Q9YFL1          Q9yfl1 aeropyrum p
 c 956    7   1.0   206   8  Q9B1X9          Q9b1x9 lutzomyia l
 c 957    7   1.0   206   8  Q9B1U7          Q9b1u7 lutzomyia l
 c 958    7   1.0   206   8  Q9B1F4          Q9b1f4 lutzomyia l
 c 959    7   1.0   206   8  Q9B1F3          Q9b1f3 lutzomyia l
 c 960    7   1.0   206   8  Q9B8C1          Q9b8c1 lutzomyia l
 c 961    7   1.0   206   8  Q9B8C0          Q9b8c0 lutzomyia l
 c 962    7   1.0   206   8  Q9B8B9          Q9b8b9 lutzomyia l
 c 963    7   1.0   206   8  Q9B8B8          Q9b8b8 lutzomyia l
 c 964    7   1.0   206   8  Q9B8B7          Q9b8b7 lutzomyia l
 c 965    7   1.0   206   8  Q9B8B6          Q9b8b6 lutzomyia l
 c 966    7   1.0   206   8  Q9B8B5          Q9b8b5 lutzomyia l
 c 967    7   1.0   206   8  Q9B8B4          Q9b8b4 lutzomyia g
 c 968    7   1.0   206  13  Q90802          Q90802 gallus gall
 c 969    7   1.0   206  16  Q98MH4          Q98mh4 rhizobium l
   970    7   1.0   207   3  Q01400          Q01400 neurospora
 c 971    7   1.0   207  10  Q9AVX6          Q9avx6 guillardia
 c 972    7   1.0   208  10  Q9XGJ5          Q9xgj5 gnetum gnem
 c 973    7   1.0   208  10  Q8S0Z1          Q8s0z1 oryza sativ
 c 974    7   1.0   209   3  O13600          O13600 schizosacch
 c 975    7   1.0   209   4  Q96SH7          Q96sh7 homo sapien
 c 976    7   1.0   209  11  Q9D9B5          Q9d9b5 mus musculu
   977    7   1.0   209  16  Q8UF91          Q8uf91 agrobacteri
   978    7   1.0   210   2  Q93HY5          Q93hy5 enterococcu
   979    7   1.0   210  10  Q01764          Q01764 phytophthor
 c 980    7   1.0   210  11  Q8R206          Q8r206 mus musculu
   981    7   1.0   211   5  Q9N4A6          Q9n4a6 caenorhabdi
   982    7   1.0   211  13  O93333          O93333 struthio ca
   983    7   1.0   211  17  Q8ZVD1          Q8zvd1 pyrobaculum
   984    7   1.0   212  15  Q9YXH2          Q9yxh2 human immun
   985    7   1.0   212  15  Q9QL60          Q9ql60 human immun
   986    7   1.0   213  11  Q9D0H4          Q9d0h4 mus musculu
   987    7   1.0   213  12  Q8V4Z0          Q8v4z0 monkeypox v
 c 988    7   1.0   214  10  Q9SC42          Q9sc42 pisum sativ
 c 989    7   1.0   215  10  Q9LRV7          Q9lrv7 arabidopsis
   990    7   1.0   215  17  O28708          O28708 archaeoglob
 c 991    7   1.0   216   2  Q8VU76          Q8vu76 lactococcus
   992    7   1.0   216   4  Q9BRA8          Q9bra8 homo sapien
 c 993    7   1.0   216   5  Q9VLD7          Q9vld7 drosophila
 c 994    7   1.0   216  10  Q41083          Q41083 pinus taeda
 c 995    7   1.0   216  10  O23595          O23595 arabidopsis
 c 996    7   1.0   216  16  O33128          O33128 mycobacteri
   997    7   1.0   217   3  Q9UTZ4          Q9utz4 schizosacch
   998    7   1.0   217   5  Q9XUA2          Q9xua2 caenorhabdi
 c 999    7   1.0   217   9  O80221          O80221 methanobact
 c1000    7   1.0   217  10  Q9LPI2          Q9lpi2 arabidopsis
```

ALIGNMENTS

```
RESULT 1
Q9GV10
ID   Q9GV10      PRELIMINARY;      PRT;   546 AA.
AC   Q9GV10;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa-related protein PoVAS1 (Fragment).
GN   POVAS1.
OS   Ephydatia fluviatilis.
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
OC   Haplosclerida; Spongillidae; Ephydatia.
OX   NCBI_TaxID=31330;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21359115; PubMed=11466525;
RA   Mochizuki K., Nishimiya-Fujisawa C., Fujisawa T.;
RT   "Universal occurrence of the vasa-related genes among metazoans and
RT   their germline expression in Hydra.";
RL   Dev. Genes Evol. 211:299-308(2001).
DR   EMBL; AB047385; BAB13310.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR002202; HMG-CoA_red.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS00318; HMG_COA_REDUCTASE_2; UNKNOWN_1.
KW   ATP-binding; Helicase.
FT   NON_TER       1      1
SQ   SEQUENCE   546 AA;  58871 MW;  6184562E9F637A40 CRC64;

Alignment Scores:
Pred. No.:              5.32e-19      Length:        546
Score:                  27.00         Matches:       27
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.73%         Indels:        0
DB:                     5             Gaps:          0

US-09-714-865-15 (1-2172) x Q9GV10 (1-546)

Qy        955 ATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAG 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 IleProIleIleLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLys 143

Qy       1015 ACTGCGGCTTTTCTCCTACCA 1035
              |||||||||||||||||||||
Db        144 ThrAlaAlaPheLeuLeuPro 150

RESULT 2
Q8QGD0
ID   Q8QGD0      PRELIMINARY;      PRT;   394 AA.
AC   Q8QGD0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DEAD box RNA helicase Vasa.
OS   Pantodon buchholtzi (Butterflyfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC   Osteoglossiformes; Pantodontidae; Pantodon.
OX   NCBI_TaxID=8276;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21906633; PubMed=11909530;
RA   Knaut H., Steinbeisser H., Schwarz H., Nusslein-Volhard C.;
RT   "An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA
RT   Translation to the Germ Cells in the Zebrafish.";
RL   Curr. Biol. 12:454-466(2002).
DR   EMBL; AF479823; AAL87142.1; -.
KW   Helicase.
SQ   SEQUENCE   394 AA;  43662 MW;  78F9ADF1A4A0A324 CRC64;

Alignment Scores:
Pred. No.:              7.47e-17      Length:        394
Score:                  25.00         Matches:       25
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.45%         Indels:        0
DB:                     13            Gaps:          0

US-09-714-865-15 (1-2172) x Q8QGD0 (1-394)

Qy       1840 AATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 AsnPheAspLeuProSerThrIleAspGluTyrValHisArgIleGlyArgThrGlyArg 305
```

```
Qy       1900 TGTGGGAATACTGGC 1914
              |||||||||||||||
Db        306 CysGlyAsnThrGly 310

RESULT 3
Q9I8L8
ID   Q9I8L8      PRELIMINARY;      PRT;   399 AA.
AC   Q9I8L8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Vasa-like protein (Fragment).
OS   Danio dangila.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=127599;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20273912; PubMed=10811828;
RA   Knaut H., Pelegri F., Bohmann K., Schwarz H., Nusslein-Volhard C.;
RT   "Zebrafish vasa RNA but not its protein is a component of the germ
RT   plasm and segregates asymmetrically before germline specification.";
RL   J. Cell Biol. 149:875-888(2000).
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AF251800; AAF74278.2; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
FT   NON_TER     399    399
SQ   SEQUENCE   399 AA;  43805 MW;  F20B0F4028FA1128 CRC64;

Alignment Scores:
Pred. No.:              8.58e-16       Length:         399
Score:                  24.00          Matches:        24
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            3.31%          Indels:         0
DB:                     13             Gaps:           0

US-09-714-865-15 (1-2172) x Q9I8L8 (1-399)

Qy       1669 GCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAG 1728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 AlaThrPheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGln 248

Qy       1729 AGAGAGCGGGAG 1740
              ||||||||||||
Db        249 ArgGluArgGlu 252

RESULT 4
Q90ZF6
ID   Q90ZF6      PRELIMINARY;      PRT;   617 AA.
AC   Q90ZF6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   VASA.
GN   OLVAS.
OS   Oryzias latipes (Medaka fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shinomiya A., Tanaka M., Kobayashi T., Nagahama Y., Hamaguchi S.;
RT   "The vasa-like gene, olvas, identifies the migration path of
RT   primordial germ cells during embryonic body formation stage in the
RT   medaka, Oryzias latipes.";
RL   Development 42:317-326(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21126920; PubMed=11226275;
RA   Tanaka M., Kinoshita M., Kobayashi D., Nagahama Y.;
RT   "Establishment of medaka (Oryzias latipes) transgenic lines with the
RT   expression of green fluorescent protein fluorescence exclusively in
RT   germ cells: A useful model to monitor germ cells in a live
RT   vertebrate.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:2544-2549(2001).
DR   EMBL; AB063484; BAB61047.1; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; UNKNOWN_1.
KW   ATP-binding; Helicase.
SQ   SEQUENCE   617 AA;  66737 MW;  C1B8A3462760BA50 CRC64;

Alignment Scores:
Pred. No.:              7.91e-16       Length:         617
Score:                  24.00          Matches:        24
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            3.31%          Indels:         0
DB:                     13             Gaps:           0

US-09-714-865-15 (1-2172) x Q90ZF6 (1-617)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 236

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        237 LeuProIleLeu 240

RESULT 5
Q8QHL6
ID   Q8QHL6      PRELIMINARY;      PRT;   621 AA.
AC   Q8QHL6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa short form.
GN   VAS-S.
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21664110; PubMed=11804791;
RA   Kobayashi T., Kajiura-Kobayashi H., Nagahama Y.;
RT   "Two isoforms of vasa homologs in a teleost fish: their differential
RT   expression during germ cell differentiation.";
RL   Mech. Dev. 111:167-171(2002).
DR   EMBL; AB051835; BAB56110.1; -.
SQ   SEQUENCE   621 AA;  67651 MW;  F0A828FDC5308FAA CRC64;

Alignment Scores:
Pred. No.:              7.9e-16        Length:         621
Score:                  24.00          Matches:        24
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
```

```
Query Match:             3.31%          Indels:        0
DB:                      13             Gaps:          0

US-09-714-865-15 (1-2172) x Q8QHL6 (1-621)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 238

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        239 LeuProIleLeu 242

RESULT 6
Q9DEG3
ID   Q9DEG3      PRELIMINARY;      PRT;   645 AA.
AC   Q9DEG3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa.
GN   VAS.
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20544961; PubMed=11091081;
RA   Kobayashi T., Kajiura-Kobayashi H., Nagahama Y.;
RT   "Differential expression of vasa homologue gene in the germ cells
RT   during oogenesis and spermatogenesis in a teleost fish, tilapia,
RT   Oreochromis niloticus.";
RL   Mech. Dev. 99:139-142(2000).
DR   EMBL; AB032467; BAB19807.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR000362; Fumarate_lyase.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000685; RuBisCO_large.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS00163; FUMARATE_LYASES; UNKNOWN_1.
DR   PROSITE; PS00157; RUBISCO_LARGE; UNKNOWN_1.
KW   ATP-binding; Helicase.
SQ   SEQUENCE   645 AA;  70072 MW;  4C945536A9BF21DA CRC64;

Alignment Scores:
Pred. No.:               7.84e-16       Length:        645
Score:                   24.00          Matches:       24
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             3.31%          Indels:        0
DB:                      13             Gaps:          0

US-09-714-865-15 (1-2172) x Q9DEG3 (1-645)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 262

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        263 LeuProIleLeu 266

RESULT 7
Q9PT10
ID   Q9PT10      PRELIMINARY;      PRT;   647 AA.
AC   Q9PT10;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update).
DE   Vasa.
GN   VAS.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=20160526; PubMed=10694742;
RA   Yoshizaki G., Sakatani S., Tominaga H., Takeuchi T.;
RT   "Cloning and characterization of a vasa-like gene in rainbow trout and
RT   its expression in the germ cell lineage.";
RL   Mol. Reprod. Dev. 55:364-371(2000).
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AB032566; BAA88059.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   647 AA;  68525 MW;  88F25AE7A68F51C2 CRC64;

Alignment Scores:
Pred. No.:               7.84e-16       Length:        647
Score:                   24.00          Matches:       24
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             3.31%          Indels:        0
DB:                      13             Gaps:          0

US-09-714-865-15 (1-2172) x Q9PT10 (1-647)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 266

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        267 LeuProIleLeu 270

RESULT 8
Q91372
ID   Q91372      PRELIMINARY;      PRT;   700 AA.
AC   Q91372;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   DEAD box protein.
GN   VLG1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94200507; PubMed=8150200;
RA   Komiya T., Itoh K., Ikenishi K., Furusawa M.;
RT   "Isolation and characterization of a novel gene of the DEAD box
RT   protein family which is specifically expressed in germ cells of
RT   Xenopus laevis.";
```

```
RL   Dev. Biol. 162:354-363(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Komiya T.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AF046043; AAC03114.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000685; RuBisCO_large.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS00157; RUBISCO_LARGE; UNKNOWN_1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   700 AA;  78240 MW;  C504ECA38EFB0B7E CRC64;

Alignment Scores:
Pred. No.:              7.72e-16      Length:         700
Score:                  24.00         Matches:        24
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            3.31%         Indels:         0
DB:                     13            Gaps:           0

US-09-714-865-15 (1-2172) x Q91372 (1-700)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 329

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        330 LeuProIleLeu 333

RESULT 9
O42378
ID   O42378      PRELIMINARY;      PRT;   700 AA.
AC   O42378;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   RNA helicase (DEAD box).
GN   VASA OR VLG.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98025484; PubMed=9376327;
RA   Olsen L.C., Aasland R., Fjose A.;
RT   "A vasa-like gene in Zebrafish identifies putative primordial germ
RT   cells.";
RL   Mech. Dev. 66:95-105(1997).
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; Y12007; CAA72735.1; -.
DR   HSSP; Q58083; 1HV8.
DR   ZFIN; ZDB-GENE-990415-272; vasa.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   700 AA;  75312 MW;  726B56A90D6C2DB3 CRC64;

Alignment Scores:
Pred. No.:              7.72e-16      Length:         700
Score:                  24.00         Matches:        24
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            3.31%         Indels:         0
DB:                     13            Gaps:           0

US-09-714-865-15 (1-2172) x O42378 (1-700)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 316

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        317 LeuProIleLeu 320

RESULT 10
Q8QGG8
ID   Q8QGG8      PRELIMINARY;      PRT;   715 AA.
AC   Q8QGG8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa-like protein.
GN   VASA.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB;
RA   Bartfai R., Orban L.;
RT   "Characterization of the genomic locus encoding vasa protein in
RT   zebrafish.";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF461759; AAL89410.1; -.
SQ   SEQUENCE   715 AA;  76795 MW;  602331B3700203AA CRC64;

Alignment Scores:
Pred. No.:              7.69e-16      Length:         715
Score:                  24.00         Matches:        24
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            3.31%         Indels:         0
DB:                     13            Gaps:           0

US-09-714-865-15 (1-2172) x Q8QGG8 (1-715)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 331

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        332 LeuProIleLeu 335

RESULT 11
Q8QFU0
ID   Q8QFU0      PRELIMINARY;      PRT;   715 AA.
AC   Q8QFU0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa-like protein.
GN   VLG.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Vatland A., Olsen L.C.;
RT   "Establishment of transgenic zebrafish lines expressing green
RT   fluorescent protein in their germ cells.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ311625; CAC84069.1; -.
SQ   SEQUENCE   715 AA;  76880 MW;  3CE4E03F02073BA4 CRC64;

Alignment Scores:
Pred. No.:              7.69e-16       Length:         715
Score:                  24.00          Matches:        24
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            3.31%          Indels:         0
DB:                     13             Gaps:           0

US-09-714-865-15 (1-2172) x Q8QFU0 (1-715)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 331

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        332 LeuProIleLeu 335

RESULT 12
O42107
ID   O42107      PRELIMINARY;      PRT;   716 AA.
AC   O42107;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Vas.
GN   VASA.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97417583; PubMed=9272956;
RA   Yoon C., Kawakami K., Hopkins N.;
RT   "Zebrafish vasa homologue RNA is localized to the cleavage planes of
RT   2-and 4-cell-stage embryos and is expressed in the primordial germ
RT   cells.";
RL   Development 124:3157-3166(1997).
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AB005147; BAA22535.1; -.
DR   HSSP; Q58083; 1HV8.
DR   ZFIN; ZDB-GENE-990415-272; vasa.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   716 AA;  77004 MW;  BB74923B95C7AC5A CRC64;

Alignment Scores:
Pred. No.:              7.69e-16       Length:         716
Score:                  24.00          Matches:        24
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            3.31%          Indels:         0
DB:                     13             Gaps:           0

US-09-714-865-15 (1-2172) x O42107 (1-716)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 332

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        333 LeuProIleLeu 336

RESULT 13
Q8QGC9
ID   Q8QGC9      PRELIMINARY;      PRT;   400 AA.
AC   Q8QGC9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DEAD box RNA helicase Vasa.
OS   Melanotaenia fluviatilis.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Atheriniformes; Melanotaenioidei; Melanotaeniidae; Melanotaenia.
OX   NCBI_TaxID=120844;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21906633; PubMed=11909530;
RA   Knaut H., Steinbeisser H., Schwarz H., Nusslein-Volhard C.;
RT   "An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA
RT   Translation to the Germ Cells in the Zebrafish.";
RL   Curr. Biol. 12:454-466(2002).
DR   EMBL; AF479824; AAL87143.1; -.
KW   Helicase.
SQ   SEQUENCE   400 AA;  43797 MW;  5BDEF82585AC9059 CRC64;

Alignment Scores:
Pred. No.:              9.86e-15       Length:         400
Score:                  23.00          Matches:        23
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            3.18%          Indels:         0
DB:                     13             Gaps:           0

US-09-714-865-15 (1-2172) x Q8QGC9 (1-400)

Qy       1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 IleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAla 312

Qy       1921 ATTTCCTTT 1929
              |||||||||
Db        313 IleSerPhe 315

RESULT 14
Q9QWS9
ID   Q9QWS9      PRELIMINARY;      PRT;   658 AA.
AC   Q9QWS9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DBY protein (DEAD (aspartate-glutamate-alanine-aspartate) box
DE   polypeptide, Y chromosome).
GN   DBY OR DBY.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=TESTIS;
RX   MEDLINE=98409500; PubMed=9736773;
```

```
RA   Mazeyrat S., Saut N., Sargent C., Grimmond S., Longepied G.,
RA   Ehrmann I.E., Ellis P.S., Greenfield A., Affara N.A., Mitchell M.J.;
RT   "The mouse Y chromosome interval necessary for spermatogonial
RT   proliferation is gene dense with syntenic homology to the human AZFa
RT   region.";
RL   Hum. Mol. Genet. 7:1713-1724(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AJ007376; CAA07483.1; -.
DR   EMBL; BC021453; AAH21453.1; -.
DR   HSSP; Q58083; 1HV8.
DR   MGD; MGI:1349406; Dby.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   658 AA;  73427 MW;  C64668326B2C3BB9 CRC64;

Alignment Scores:
Pred. No.:              1.03e-13      Length:        658
Score:                  22.00         Matches:       22
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.04%         Indels:        0
DB:                     11            Gaps:          0

US-09-714-865-15 (1-2172) x Q9QWS9 (1-658)

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 236

Qy       1036 ATTTTG 1041
              ||||||
Db        237 IleLeu 238

RESULT 15
Q8QGC8
ID   Q8QGC8      PRELIMINARY;      PRT;   396 AA.
AC   Q8QGC8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DEAD box RNA helicase Vasa.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21906633; PubMed=11909530;
RA   Knaut H., Steinbeisser H., Schwarz H., Nusslein-Volhard C.;
RT   "An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA
RT   Translation to the Germ Cells in the Zebrafish.";
RL   Curr. Biol. 12:454-466(2002).
DR   EMBL; AF479825; AAL87144.1; -.
KW   Helicase.
SQ   SEQUENCE   396 AA;  43087 MW;  FFDF9B3FDED17BA0 CRC64;

Alignment Scores:
Pred. No.:              1.31e-12      Length:        396
Score:                  21.00         Matches:       21
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.90%         Indels:        0
DB:                     13            Gaps:          0

US-09-714-865-15 (1-2172) x Q8QGC8 (1-396)

Qy       1699 ACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTT 1758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 ThrThrSerIleHisGlyAspArgGluGlnArgGluArgGluGlnAlaLeuGlyAspPhe 258

Qy       1759 CGC 1761
              |||
Db        259 Arg 259

Search completed: June 10, 2003, 17:08:43
Job time : 115 secs
```

OM nucleic - protein search, using frame_plus_n2p model

```
Run on:         June 10, 2003, 16:45:36 ; Search time 21.5 Seconds
                                          (without alignments)
                                          8380.147 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  724
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  OLIGO
                Xgapop 60.0 , Xgapext 60.0
                Ygapop 60.0 , Ygapext 60.0
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       112892 seqs, 41476328 residues

Word size:      1

Total number of hits satisfying chosen parameters:      225614

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111832_26571/app_query.fasta_1.2311
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=oligo.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=1000 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_23_@runat_05062003_111832_26571 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :      SwissProt_40:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	611	84.4	724	1	DDX4_HUMAN	Q9nqi0 homo sapien
	2	82	11.3	702	1	DDX4_MOUSE	Q61496 mus musculu
	3	82	11.3	713	1	DDX4_RAT	Q64060 rattus norv
	4	23	3.2	661	1	VASA_DROME	P09052 drosophila
	5	22	3.0	660	1	PL10_MOUSE	P16381 mus musculu
	6	22	3.0	661	1	DDX3_HUMAN	O00571 homo sapien
	7	22	3.0	661	1	DDX3_MOUSE	Q62167 mus musculu
	8	22	3.0	697	1	AN3_XENLA	P24346 xenopus lae
	9	18	2.5	660	1	DDXY_HUMAN	O15523 homo sapien
	10	18	2.5	763	1	GLH1_CAEEL	P34689 caenorhabdi
	11	17	2.3	617	1	DBP1_YEAST	P24784 saccharomyc
	12	17	2.3	720	1	GLH3_CAEEL	O01836 caenorhabdi
	13	16	2.2	245	1	DEA2_MOUSE	Q62095 mus musculu
	14	16	2.2	604	1	DED1_YEAST	P06634 saccharomyc
	15	16	2.2	636	1	DED1_SCHPO	O13370 schizosacch
	16	15	2.1	546	1	DBP2_YEAST	P24783 saccharomyc
	17	15	2.1	550	1	DBP2_SCHPO	P24782 schizosacch
	18	15	2.1	568	1	HE64_TRYBB	Q26696 trypanosoma
	19	15	2.1	575	1	RM62_DROME	P19109 drosophila
	20	15	2.1	607	1	DB10_NICSY	P46942 nicotiana s
	21	15	2.1	614	1	DDX5_HUMAN	P17844 homo sapien
	22	15	2.1	614	1	DDX5_MOUSE	Q61656 mus musculu
	23	15	2.1	650	1	DD17_HUMAN	Q92841 homo sapien
	24	14	1.9	454	1	RHLE_ECOLI	P25888 escherichia
	25	12	1.7	444	1	SRMB_ECOLI	P21507 escherichia
	26	12	1.7	578	1	YBX6_SCHPO	Q10202 schizosacch
	27	11	1.5	457	1	DBPA_ECOLI	P21693 escherichia
	28	11	1.5	523	1	DBP3_YEAST	P20447 saccharomyc
	29	11	1.5	543	1	RRP3_YEAST	P38712 saccharomyc
	30	11	1.5	601	1	DEAD_BUCAI	P57453 buchnera ap
	31	11	1.5	613	1	DEAD_HAEIN	P44586 haemophilus
	32	11	1.5	628	1	DEAD_ECOLI	P23304 escherichia
	33	11	1.5	642	1	DEAD_KLEPN	P33906 klebsiella
	34	11	1.5	1156	1	GLH4_CAEEL	O76743 caenorhabdi
	35	10	1.4	418	1	RHLB_HAEIN	P44922 haemophilus
	36	10	1.4	420	1	RHLB_ECOLI	P24229 escherichia
	37	10	1.4	420	1	RHLB_SALTY	P40863 salmonella
	38	10	1.4	524	1	EXP9_STRPN	P35599 streptococc
	39	10	1.4	563	1	DEAD_MYCTU	Q11039 mycobacteri
	40	10	1.4	564	1	ROK1_YEAST	P45818 saccharomyc
	41	10	1.4	619	1	ABS_DROME	Q9v3c0 drosophila
	42	10	1.4	622	1	ABS_HUMAN	Q9ujv9 homo sapien
c	43	9	1.3	168	1	Y168_ADE02	P03292 human adeno
	44	9	1.2	431	1	DBP8_YEAST	P38719 saccharomyc
c	45	9	1.3	437	1	TIS1_DROME	P47980 drosophila
	46	9	1.2	439	1	DDX7_HUMAN	Q15320 homo sapien
	47	9	1.2	439	1	SRMB_HAEIN	P44701 haemophilus
c	48	9	1.3	453	1	EZY1_CHLRE	Q08354 chlamydomon
	49	9	1.2	481	1	YA88_SCHPO	Q09775 schizosacch
c	50	9	1.3	508	1	EGR1_RAT	P08154 rattus norv
c	51	9	1.3	543	1	EGR1_HUMAN	P18146 homo sapien
c	52	9	1.3	579	1	ESR1_SPAAU	Q9pvz9 sparus aura
	53	9	1.2	606	1	SPB4_YEAST	P25808 saccharomyc
c	54	9	1.3	663	1	DUS8_MOUSE	O09112 mus musculu
	55	9	1.2	752	1	DRS1_YEAST	P32892 saccharomyc
	56	9	1.2	754	1	YAJ3_SCHPO	Q09903 schizosacch
c	57	9	1.3	788	1	SP1_RAT	Q01714 rattus norv
	58	9	1.2	848	1	YA47_SCHPO	Q09719 schizosacch
c	59	9	1.3	958	1	TRH_DROME	Q24119 drosophila
	60	9	1.2	995	1	DBPA_YEAST	Q12389 saccharomyc
	61	9	1.2	1193	1	DP3A_XYLFA	Q9pgu4 xylella fas
c	62	8	1.1	115	1	KV2A_HUMAN	P01614 homo sapien
	63	8	1.1	143	1	RNH_HELPY	P56120 helicobacte
c	64	8	1.1	265	1	UCRI_SOLTU	P37841 solanum tub
c	65	8	1.1	267	1	COLI_PIG	P01192 s corticotr
c	66	8	1.1	271	1	PDX4_HUMAN	Q13162 homo sapien
c	67	8	1.1	274	1	PDX4_BOVIN	Q9bgi2 bos taurus
c	68	8	1.1	274	1	PDX4_MOUSE	O08807 mus musculu
c	69	8	1.1	297	1	EZY3_CHLRE	Q08356 chlamydomon
c	70	8	1.1	320	1	CH36_CERCA	P17110 ceratitis c
	71	8	1.1	327	1	MALR_STRPN	Q08511 streptococc
	72	8	1.1	367	1	H669_METJA	Q58083 methanococc
c	73	8	1.1	367	1	HK61_HUMAN	P78426 homo sapien
c	74	8	1.1	416	1	NAP5_HUMAN	O14513 homo sapien
	75	8	1.1	460	1	DB80_DROME	O61305 drosophila
c	76	8	1.1	464	1	IFE_BRALA	Q04948 branchiosto
c	77	8	1.1	469	1	DPD2_BOVIN	P49004 bos taurus
	78	8	1.1	478	1	DD19_MOUSE	Q61655 mus musculu
	79	8	1.1	479	1	DBPA_BACSU	P42305 bacillus su
	80	8	1.1	479	1	DD19_HUMAN	Q9umr2 homo sapien
	81	8	1.1	482	1	DBP5_YEAST	P20449 saccharomyc
	82	8	1.1	489	1	YN21_CAEEL	P34580 caenorhabdi
	83	8	1.1	503	1	YB66_SCHPO	Q09747 schizosacch
	84	8	1.1	505	1	HAS1_YEAST	Q03532 saccharomyc
c	85	8	1.1	518	1	FUS_MOUSE	P56959 mus musculu
	86	8	1.1	521	1	DB45_DROME	Q07886 drosophila
c	87	8	1.1	526	1	FUS_HUMAN	P35637 homo sapien
	88	8	1.1	546	1	NOH6_MOUSE	Q9d0r4 mus musculu
	89	8	1.1	547	1	NOH6_HUMAN	Q9ny93 homo sapien
	90	8	1.1	564	1	ZYX_MOUSE	Q62523 mus musculu
c	91	8	1.1	569	1	K1CJ_MOUSE	P02535 mus musculu

	92	8	1.1	572	1	ZYX_HUMAN	Q15942	homo sapien
	93	8	1.1	578	1	YAK2_SCHPO	Q09916	schizosacch
c	94	8	1.1	585	1	ESR1_ORENI	Q9yh33	oreochromis
	95	8	1.1	588	1	PR28_YEAST	P23394	saccharomyc
c	96	8	1.1	617	1	41_BOVIN	Q9n179	bos taurus
	97	8	1.1	633	1	NAH2_YEAST	Q04121	saccharomyc
	98	8	1.1	663	1	PIT_DROME	Q9vd51	drosophila
	99	8	1.1	670	1	DD18_HUMAN	Q9nvp1	homo sapien
	100	8	1.1	715	1	DD21_HUMAN	Q9nr30	homo sapien
	101	8	1.1	742	1	DBP7_YEAST	P36120	saccharomyc
	102	8	1.1	752	1	PHSM_STRPN	P29849	streptococc
	103	8	1.1	770	1	DBP4_YEAST	P20448	saccharomyc
	104	8	1.1	773	1	MAK5_YEAST	P38112	saccharomyc
c	105	8	1.1	801	1	41_XENLA	P11434	xenopus lae
c	106	8	1.1	805	1	E2F_DROME	Q27368	drosophila
c	107	8	1.1	823	1	NUC1_NEUCR	P20824	neurospora
	108	8	1.1	851	1	DD21_MOUSE	Q9jik5	mus musculu
c	109	8	1.1	854	1	VG12_BPPZA	P07537	bacteriopha
	110	8	1.1	857	1	DD24_MOUSE	Q9esv0	mus musculu
c	111	8	1.1	858	1	41_MOUSE	P48193	mus musculu
	112	8	1.1	859	1	DD24_HUMAN	Q9gzr7	homo sapien
c	113	8	1.1	863	1	SYV_ARCFU	O28059	archaeoglob
	114	8	1.1	875	1	DD10_HUMAN	Q13206	homo sapien
	115	8	1.1	971	1	SEC5_YEAST	P89102	saccharomyc
c	116	8	1.1	1043	1	CHS2_PARBR	Q92444	paracoccidi
	117	8	1.1	1225	1	NFT5_MOUSE	Q9wv30	mus musculu
c	118	8	1.1	1295	1	GLP1_CAEEL	P13508	caenorhabdi
c	119	8	1.1	1324	1	SAL1_HUMAN	Q9nsc2	homo sapien
	120	8	1.1	1531	1	NFT5_HUMAN	O94916	homo sapien
c	121	8	1.1	2142	1	BAT2_HUMAN	P48634	homo sapien
	122	8	1.1	2175	1	HMCU_DROME	P10180	drosophila
	123	7	1.0	52	1	ATP8_PENMO	Q9mgd7	penaeus mon
c	124	7	1.0	64	1	RL28_MYCGE	P47665	mycoplasma
c	125	7	1.0	64	1	RL35_MYCLE	Q9cc21	mycobacteri
c	126	7	1.0	73	1	YC40_ODOSI	P49535	odontella s
	127	7	1.0	91	1	RS7_APPPP	Q9zeu5	apple proli
c	128	7	1.0	91	1	VE5_HPV13	Q02267	human papil
c	129	7	1.0	96	1	CLV3_ARATH	Q9xf04	arabidopsis
	130	7	1.0	104	1	YI64_ARCFU	O28415	archaeoglob
c	131	7	1.0	110	1	GON2_SUNMU	O97686	suncus muri
	132	7	1.0	112	1	SY27_HUMAN	Q9y4x3	homo sapien
c	133	7	1.0	114	1	GON2_TUPGB	Q95336	tupaia glis
	134	7	1.0	116	1	YJ57_YEAST	P47132	saccharomyc
c	135	7	1.0	117	1	NU3M_PROWI	Q37625	prototheca
c	136	7	1.0	120	1	GON2_HUMAN	O43555	homo sapien
	137	7	1.0	120	1	R18E_THEAC	Q9hl06	thermoplasm
	138	7	1.0	132	1	FOLB_MYCLE	O69529	mycobacteri
c	139	7	1.0	133	1	POP8_YEAST	P38208	saccharomyc
c	140	7	1.0	133	1	RS9_DEIRA	Q9rxy0	deinococcus
c	141	7	1.0	147	1	RNL4_HUMAN	P34096	homo sapien
c	142	7	1.0	147	1	RNL4_PIG	P15468	sus scrofa
c	143	7	1.0	147	1	RNL4_RAT	O55004	rattus norv
c	144	7	1.0	148	1	RNL4_MOUSE	Q9jjh1	mus musculu
	145	7	1.0	150	1	AROQ_BUCAI	P57479	buchnera ap
	146	7	1.0	154	1	SM20_SCHMA	P15845	schistosoma
	147	7	1.0	155	1	RS7_MYCLE	P30764	mycobacteri
	148	7	1.0	155	1	RS7_MYCSM	P41193	mycobacteri
	149	7	1.0	155	1	RS7_MYCTU	P41194	mycobacteri
c	150	7	1.0	155	1	Y805_AQUAE	O66989	aquifex aeo
	151	7	1.0	156	1	RS7_MICLU	P09898	micrococcus
	152	7	1.0	157	1	RS7_CHLMU	Q9pjv5	chlamydia m
	153	7	1.0	157	1	RS7_CHLPN	Q9z801	chlamydia p
	154	7	1.0	157	1	RS7_CHLTR	P29765	chlamydia t
c	155	7	1.0	161	1	UCN3_HUMAN	Q969e3	homo sapien
	156	7	1.0	162	1	CRF_XENLA	P49188	xenopus lae
	157	7	1.0	163	1	YCBL_BACUN	P30906	bacteroides
	158	7	1.0	164	1	YJJA_ECOLI	P18390	escherichia
	159	7	1.0	165	1	NU6M_STRPU	P15553	strongyloce
	160	7	1.0	167	1	NU6M_MYXGL	Q9g2w7	myxine glut
c	161	7	1.0	171	1	ATPF_SYNP6	P08447	synechococc
c	162	7	1.0	176	1	TR23_MOUSE	Q9er63	mus musculu
	163	7	1.0	178	1	RR7_ASTLO	P14760	astasia lon
	164	7	1.0	182	1	FTN_DROME	P35554	drosophila
	165	7	1.0	185	1	YWLD_BACSU	P39154	bacillus su
	166	7	1.0	186	1	DNAA_WOLSP	P35907	wolbachia s
	167	7	1.0	189	1	YD48_AQUAE	O67363	aquifex aeo
c	168	7	1.0	190	1	BCT7_SHEEP	P50415	ovis aries
c	169	7	1.0	192	1	PABA_STRLI	P27627	streptomyce
c	170	7	1.0	192	1	PTH_AQUAE	O66677	aquifex aeo
c	171	7	1.0	194	1	AX2A_PHAAU	P32293	phaseolus a
	172	7	1.0	194	1	VATE_HALVO	Q48329	halobacteri
c	173	7	1.0	202	1	YIBF_ECOLI	P32105	escherichia
c	174	7	1.0	205	1	YKO7_YEAST	P36061	saccharomyc
c	175	7	1.0	206	1	UCRI_CHLRE	P49728	chlamydomon
	176	7	1.0	209	1	YQIJ_ECOLI	P76657	escherichia
	177	7	1.0	210	1	YG2I_YEAST	P53251	saccharomyc
	178	7	1.0	221	1	IER2_MOUSE	P17950	mus musculu
	179	7	1.0	223	1	TONB_XANCP	O34261	xanthomonas
	180	7	1.0	228	1	UREF_YEREN	P42870	yersinia en
	181	7	1.0	228	1	UREF_YERPE	Q9zfr7	yersinia pe
	182	7	1.0	228	1	UREF_YERPS	P52318	yersinia ps
	183	7	1.0	228	1	VIF_OMVVS	P16902	ovine lenti
c	184	7	1.0	229	1	RAS2_NEUCR	Q01387	neurospora
	185	7	1.0	229	1	VIF_CAEVC	P33462	caprine art
	186	7	1.0	230	1	VIF_VILV	P03403	visna lenti
	187	7	1.0	230	1	VIF_VILV1	P23430	visna lenti
c	188	7	1.0	232	1	YM29_MYCLE	Q9cbs9	mycobacteri
	189	7	1.0	236	1	TL17_ARATH	P81760	arabidopsis
c	190	7	1.0	237	1	ATP6_PETMA	Q35538	petromyzon
	191	7	1.0	238	1	ATPI_CHLRE	O63075	chlamydomon
c	192	7	1.0	239	1	GIDB_CLOAB	Q97cw4	clostridium
c	193	7	1.0	240	1	DLX4_MOUSE	P70436	mus musculu
	194	7	1.0	243	1	PLT_HUMAN	Q02083	homo sapien
	195	7	1.0	247	1	ATPI_CHLVU	P56295	chlorella v
c	196	7	1.0	247	1	MCT1_HUMAN	P23946	homo sapien
c	197	7	1.0	247	1	MCT1_MACFA	P56435	macaca fasc
c	198	7	1.0	247	1	MCT1_PAPHA	P52195	papio hamad
c	199	7	1.0	247	1	MOG_HUMAN	Q16653	homo sapien
c	200	7	1.0	248	1	GRL1_RAT	Q06605	rattus norv
	201	7	1.0	250	1	ETRA_SHEON	P46148	shewanella
c	202	7	1.0	251	1	YMV5_CAEEL	P34507	caenorhabdi
c	203	7	1.0	253	1	LEP4_VIBCH	P27717	vibrio chol
	204	7	1.0	254	1	UL79_HSV7J	P52470	human herpe
c	205	7	1.0	256	1	PXBA_CANBO	Q00316	candida boi
c	206	7	1.0	256	1	PXBB_CANBO	Q00317	candida boi
	207	7	1.0	257	1	ECH8_MYCLE	O07137	mycobacteri
	208	7	1.0	257	1	NT6A_HUMAN	P34132	homo sapien
	209	7	1.0	257	1	NT6B_HUMAN	P34133	homo sapien
	210	7	1.0	258	1	YC10_METJA	Q58607	methanococc
c	211	7	1.0	259	1	EMD_MOUSE	O08579	mus musculu
c	212	7	1.0	259	1	MSRA_LACSA	Q9sec2	lactuca sat
c	213	7	1.0	262	1	GDA1_WHEAT	P04721	triticum ae
c	214	7	1.0	266	1	ARC4_PHAVU	Q43629	phaseolus v
c	215	7	1.0	267	1	MM07_RAT	P50280	rattus norv
c	216	7	1.0	268	1	EP34_HCMVA	P16768	human cytom
c	217	7	1.0	273	1	YAFZ_ECOLI	P77206	escherichia
c	218	7	1.0	273	1	YFJQ_ECOLI	P52132	escherichia
c	219	7	1.0	274	1	GLPF_BACSU	P18156	bacillus su
	220	7	1.0	274	1	STCQ_EMENI	Q00713	emericella
	221	7	1.0	279	1	UVEN_MICLU	P46303	micrococcus
c	222	7	1.0	286	1	CYSW_SYNP7	P27370	synechococc
c	223	7	1.0	286	1	GDAO_WHEAT	P02863	triticum ae
c	224	7	1.0	294	1	CHIB_TOBAC	P29061	nicotiana t
c	225	7	1.0	296	1	GDA6_WHEAT	P04726	triticum ae
c	226	7	1.0	296	1	SAPC_ECOLI	Q47624	escherichia
c	227	7	1.0	296	1	SAPC_SALTY	P36669	salmonella
	228	7	1.0	296	1	THTM_RAT	P97532	rattus norv
c	229	7	1.0	297	1	GDA4_WHEAT	P04724	triticum ae
c	230	7	1.0	301	1	HXDD_CHICK	P24344	gallus gall
c	231	7	1.0	303	1	GDFF_MOUSE	Q9z0j7	mus musculu
c	232	7	1.0	303	1	GDFF_RAT	Q9z0j6	rattus norv
c	233	7	1.0	304	1	CBPZ_SIMVI	P42788	simulium vi
	234	7	1.0	304	1	PEX2_CRIGR	Q06438	cricetulus
	235	7	1.0	305	1	PEX2_HUMAN	P28328	homo sapien
	236	7	1.0	305	1	PEX2_MOUSE	P55098	mus musculu
	237	7	1.0	305	1	YC44_LISIN	Q925x2	listeria in

c	238	7	1.0	307	1	GDA9_WHEAT	P18573 triticum ae
c	239	7	1.0	313	1	GDA7_WHEAT	P04727 triticum ae
	240	7	1.0	314	1	MIAA_CHLMU	Q9plf7 chlamydia m
	241	7	1.0	314	1	MIAA_CHLTR	O84771 chlamydia t
c	242	7	1.0	314	1	NMI_MOUSE	O35309 mus musculu
c	243	7	1.0	316	1	LORI_HUMAN	P23490 homo sapien
c	244	7	1.0	319	1	GDA5_WHEAT	P04725 triticum ae
	245	7	1.0	324	1	YL70_ARCFU	O28112 archaeoglob
c	246	7	1.0	325	1	TRBB_RHISN	P55395 rhizobium s
c	247	7	1.0	329	1	YG54_ARCFU	O28619 archaeoglob
	248	7	1.0	334	1	WDR5_HUMAN	Q9ugp9 homo sapien
c	249	7	1.0	335	1	FLIG_THEMA	Q9wy63 thermotoga
c	250	7	1.0	337	1	FSA_SHEEP	P31514 ovis aries
c	251	7	1.0	337	1	TAL1_MOUSE	Q93092 mus musculu
c	252	7	1.0	341	1	FSA_XENLA	P31515 xenopus lae
c	253	7	1.0	343	1	FSA_CHICK	Q90844 gallus gall
c	254	7	1.0	344	1	FSA_BOVIN	P50291 bos taurus
c	255	7	1.0	344	1	FSA_HORSE	O62650 equus cabal
c	256	7	1.0	344	1	FSA_MOUSE	P47931 mus musculu
c	257	7	1.0	344	1	FSA_PIG	P10669 sus scrofa
c	258	7	1.0	344	1	FSA_RAT	P21674 rattus norv
c	259	7	1.0	345	1	YLX4_CAEEL	P46500 caenorhabdi
	260	7	1.0	348	1	CHLI_EUGGR	P31205 euglena gra
c	261	7	1.0	348	1	N33_HUMAN	Q13454 homo sapien
	262	7	1.0	351	1	HI81_PASMU	Q9clm3 pasteurella
c	263	7	1.0	355	1	HKL6_LYCES	O22299 lycopersico
c	264	7	1.0	355	1	LAV1_PHYPO	P14725 physarum po
c	265	7	1.0	356	1	SMR1_PODAN	Q08142 podospora a
	266	7	1.0	357	1	RUVB_STRCO	Q91291 streptomyce
c	267	7	1.0	360	1	GP25_HUMAN	O00155 homo sapien
c	268	7	1.0	361	1	ELYA_BACHD	P41363 bacillus ha
	269	7	1.0	361	1	WDS_DROME	Q9v3j8 drosophila
	270	7	1.0	361	1	YIDS_ECO57	Q8xbz8 escherichia
c	271	7	1.0	366	1	IHA_MOUSE	Q04997 mus musculu
c	272	7	1.0	366	1	IHA_RAT	P17490 rattus norv
c	273	7	1.0	372	1	3BH1_MOUSE	P24815 m 3 beta-hy
c	274	7	1.0	372	1	3BH1_RAT	P22071 r 3 beta-hy
c	275	7	1.0	372	1	3BH2_RAT	P22072 r 3 beta-hy
c	276	7	1.0	372	1	3BH3_RAT	P27364 rattus norv
c	277	7	1.0	372	1	3BH4_MOUSE	Q61767 mus musculu
c	278	7	1.0	372	1	3BH4_RAT	Q62878 r 3 beta-hy
c	279	7	1.0	373	1	SRB3_HUMAN	Q9ns66 homo sapien
c	280	7	1.0	379	1	CIT2_ECOLI	P05853 escherichia
c	281	7	1.0	380	1	GAT4_CHICK	P43691 gallus gall
	282	7	1.0	380	1	TRPD_YEAST	P07285 saccharomyc
c	283	7	1.0	381	1	CYB_NOTTY	O03478 notoryctes
c	284	7	1.0	383	1	NAPA_ENTHR	P26235 enterococcu
c	285	7	1.0	389	1	SERI_BOMMO	P07856 bombyx mori
c	286	7	1.0	389	1	WN2B_MOUSE	O70283 mus musculu
c	287	7	1.0	390	1	PER_DROTP	P91716 drosophila
c	288	7	1.0	391	1	CAR2_RHINI	P43231 rhizopus ni
c	289	7	1.0	391	1	WN2B_HUMAN	Q93097 homo sapien
c	290	7	1.0	392	1	TRE1_DROME	Q9ndm2 drosophila
c	291	7	1.0	392	1	Y462_TREPA	O83475 treponema p
c	292	7	1.0	394	1	ACKA_LACSK	Q9x4m1 lactobacill
	293	7	1.0	396	1	CXA7_HUMAN	P36383 homo sapien
c	294	7	1.0	397	1	HKL2_MALDO	O04135 malus domes
	295	7	1.0	397	1	LIPG_BOVIN	Q29458 bos taurus
c	296	7	1.0	398	1	HKL1_MALDO	O04134 malus domes
c	297	7	1.0	398	1	Y306_CHLMU	Q9pl02 chlamydia m
	298	7	1.0	399	1	FAL1_YEAST	Q12099 saccharomyc
c	299	7	1.0	401	1	UXUA_BRUME	Q8ycq4 brucella me
	300	7	1.0	403	1	IF4A_LEIBR	Q25225 leishmania
c	301	7	1.0	403	1	SHBG_RAT	P08689 rattus norv
c	302	7	1.0	404	1	CAZ_DROME	Q27294 drosophila
c	303	7	1.0	404	1	YS69_MYCTU	O33351 mycobacteri
	304	7	1.0	405	1	ACKA_BUCAI	P57272 buchnera ap
	305	7	1.0	411	1	IF4N_HUMAN	P38919 homo sapien
c	306	7	1.0	411	1	IHH_MOUSE	P97812 mus musculu
c	307	7	1.0	412	1	DJA2_HUMAN	O60884 homo sapien
c	308	7	1.0	412	1	DJA2_MOUSE	Q9qyj0 mus musculu
c	309	7	1.0	412	1	DJA2_RAT	O35824 rattus norv
	310	7	1.0	412	1	IF41_ARATH	P41376 arabidopsis

	311	7	1.0	412	1	IF42_ARATH	P41377 arabidopsis
c	312	7	1.0	412	1	YAG7_SCHPO	Q09871 schizosacch
	313	7	1.0	413	1	IF4Z_TOBAC	Q40468 nicotiana t
c	314	7	1.0	414	1	TYY1_HUMAN	P25490 homo sapien
c	315	7	1.0	414	1	YANA_SCHPO	Q10075 schizosacch
c	316	7	1.0	416	1	SO_DROME	Q27350 drosophila
c	317	7	1.0	420	1	PAR1_XENLA	P47749 xenopus lae
	318	7	1.0	423	1	SYH_PASMU	P57988 pasteurella
	319	7	1.0	427	1	YDFJ_ECOLI	P77228 escherichia
	320	7	1.0	433	1	ENOA_MOUSE	P17182 mus musculu
	321	7	1.0	433	1	ENOA_RAT	P04764 rattus norv
	322	7	1.0	433	1	ENO_DROME	P15007 drosophila
c	323	7	1.0	433	1	HXB3_MOUSE	P09026 mus musculu
c	324	7	1.0	433	1	NU4M_RHISA	O99825 rhipicephal
c	325	7	1.0	433	1	VPAP_HCMVA	P16790 human cytom
c	326	7	1.0	436	1	CCKR_MOUSE	O08786 mus musculu
	327	7	1.0	436	1	ENO1_ENTHI	P51555 entamoeba h
	328	7	1.0	436	1	ENO_NEOFR	P42894 neocallimas
c	329	7	1.0	437	1	NU4M_ALBCO	P48914 albinaria c
c	330	7	1.0	443	1	NU4M_CHLRE	P20113 chlamydomon
	331	7	1.0	445	1	ENO1_HEVBR	Q9lej0 hevea brasi
	332	7	1.0	445	1	ENO2_HEVBR	Q9lei9 hevea brasi
	333	7	1.0	445	1	ENO_RICCO	P42896 ricinus com
c	334	7	1.0	445	1	TEA2_MOUSE	P48301 mus musculu
c	335	7	1.0	449	1	C13A_MYCTU	O08447 mycobacteri
	336	7	1.0	449	1	Y425_MYCGE	P47664 mycoplasma
	337	7	1.0	450	1	Y425_MYCPN	P75172 mycoplasma
c	338	7	1.0	452	1	NU4M_BRAFL	O47423 branchiosto
c	339	7	1.0	452	1	NU4M_BRALA	O79421 branchiosto
c	340	7	1.0	452	1	YK97_MYCLE	P54077 mycobacteri
c	341	7	1.0	452	1	YK97_MYCTU	Q10706 mycobacteri
c	342	7	1.0	453	1	AS16_HUMAN	Q96ns5 homo sapien
	343	7	1.0	458	1	ME31_DROME	P23128 drosophila
	344	7	1.0	458	1	YMV6_CAEEL	P34508 caenorhabdi
c	345	7	1.0	459	1	NU4M_PELSU	O79677 pelomedusa
c	346	7	1.0	465	1	VP19_HSVEB	P28935 equine herp
c	347	7	1.0	466	1	GUN5_THEFU	Q01786 thermomonos
c	348	7	1.0	467	1	C761_SOLME	P37121 solanum mel
c	349	7	1.0	469	1	DPD2_HUMAN	P49005 homo sapien
c	350	7	1.0	470	1	PROP_CAVPO	Q64181 cavia porce
c	351	7	1.0	474	1	DLDH_HALVO	Q04829 halobacteri
c	352	7	1.0	480	1	PRTP_HUMAN	P10619 homo sapien
c	353	7	1.0	481	1	LORI_MOUSE	P18165 mus musculu
c	354	7	1.0	482	1	CATA_BORPE	P48062 bordetella
c	355	7	1.0	482	1	K6B2_HUMAN	Q9ubs0 h ribosomal
	356	7	1.0	482	1	TPR_PORGI	P25806 porphyromon
c	357	7	1.0	485	1	GLR_RAT	P30082 rattus norv
c	358	7	1.0	487	1	AMPL_MOUSE	Q9cpy7 mus musculu
	359	7	1.0	489	1	NIFD_THIFE	P06662 thiobacillu
	360	7	1.0	491	1	NAR1_YEAST	P23503 saccharomyc
c	361	7	1.0	492	1	MM11_MOUSE	Q02853 mus musculu
c	362	7	1.0	497	1	HMES_DROME	P18488 drosophila
c	363	7	1.0	497	1	NU4M_ACACA	Q37375 acanthamoeb
c	364	7	1.0	500	1	PTGI_BOVIN	Q29626 bos taurus
c	365	7	1.0	501	1	C72N_ARATH	Q9ltm0 arabidopsis
	366	7	1.0	501	1	CSD2_DROME	Q9vmt6 drosophila
c	367	7	1.0	504	1	ETS4_DROME	P29775 drosophila
	368	7	1.0	506	1	VE2_HPV47	P22420 human papil
c	369	7	1.0	511	1	RT04_PROWI	P46743 prototheca
c	370	7	1.0	512	1	FUS_BOVIN	Q28009 bos taurus
	371	7	1.0	513	1	C942_VICSA	P98188 vicia sativ
	372	7	1.0	513	1	CDP2_MAIZE	P49101 zea mays (m
c	373	7	1.0	517	1	MTC1_COREQ	P42828 corynebacte
c	374	7	1.0	518	1	COX1_POLOR	Q95911 polypterus
c	375	7	1.0	518	1	YHG9_YEAST	P38758 saccharomyc
	376	7	1.0	520	1	RAG2_XENLA	Q91830 xenopus lae
c	377	7	1.0	521	1	NPRE_BACSU	P06142 bacillus su
c	378	7	1.0	525	1	ESR1_MICUN	P57753 micropogoni
c	379	7	1.0	526	1	HYFF_ECOLI	P77437 escherichia
c	380	7	1.0	526	1	K1CJ_BOVIN	P06394 bos taurus
c	381	7	1.0	527	1	DPE2_HUMAN	P56282 homo sapien
c	382	7	1.0	527	1	NU5M_CAEEL	P24896 caenorhabdi
	383	7	1.0	528	1	RAG2_CHICK	P25022 gallus gall

c	384	7	1.0	529	1	YB89_YEAST	P38140	saccharomyc
c	385	7	1.0	532	1	BTB3_MOUSE	P58545	mus musculu
	386	7	1.0	532	1	CDPK_DAUCA	P28582	daucus caro
	387	7	1.0	533	1	CDP2_ORYSA	P53683	oryza sativ
c	388	7	1.0	533	1	EGR1_MOUSE	P08046	mus musculu
	389	7	1.0	534	1	CDP1_ORYSA	P53682	oryza sativ
c	390	7	1.0	538	1	GLE1_YEAST	Q12315	saccharomyc
	391	7	1.0	539	1	DOP2_DROME	Q24563	drosophila
	392	7	1.0	541	1	NU57_YEAST	P48837	saccharomyc
	393	7	1.0	542	1	CDP3_ORYSA	P53684	oryza sativ
	394	7	1.0	543	1	TCPH_HUMAN	Q99832	homo sapien
	395	7	1.0	544	1	TCPH_MOUSE	P80313	mus musculu
	396	7	1.0	547	1	YMX1_CAEEL	P34509	caenorhabdi
c	397	7	1.0	549	1	DSX_DROME	P23023	drosophila
c	398	7	1.0	553	1	FXC1_HUMAN	Q12948	homo sapien
c	399	7	1.0	553	1	FXC1_MOUSE	Q61572	mus musculu
c	400	7	1.0	553	1	TF3B_CANAL	P43072	candida alb
	401	7	1.0	554	1	KLP3_SCHPO	Q9us60	schizosacch
	402	7	1.0	557	1	G6PI_KLULA	P12341	kluyveromyc
c	403	7	1.0	564	1	VA55_VACCC	P21073	vaccinia vi
c	404	7	1.0	564	1	VA55_VACCV	P24768	vaccinia vi
c	405	7	1.0	567	1	CHI3_CANAL	P40954	candida alb
	406	7	1.0	572	1	SYK_CAEEL	Q22099	caenorhabdi
	407	7	1.0	572	1	YO62_DEIRA	Q9rrm6	deinococcus
c	408	7	1.0	573	1	AMH2_HUMAN	Q16671	homo sapien
	409	7	1.0	577	1	CHVG_RHIME	P72292	rhizobium m
c	410	7	1.0	578	1	FD21_MYCTU	O50441	mycobacteri
c	411	7	1.0	579	1	FD21_MYCLE	P54200	mycobacteri
c	412	7	1.0	580	1	YML1_ARATH	O22752	arabidopsis
c	413	7	1.0	581	1	ESR1_PAGMA	O42132	pagrus majo
	414	7	1.0	585	1	MPP3_HUMAN	Q13368	homo sapien
	415	7	1.0	587	1	NDC2_RAT	P70545	rattus norv
	416	7	1.0	590	1	PHO4_NEUCR	P15710	neurospora
	417	7	1.0	590	1	SYKC_YEAST	P15180	saccharomyc
c	418	7	1.0	593	1	K1CJ_HUMAN	P13645	homo sapien
c	419	7	1.0	594	1	DBP9_YEAST	Q06218	saccharomyc
c	420	7	1.0	606	1	MM17_HUMAN	Q9ulz9	homo sapien
	421	7	1.0	614	1	ZF29_MOUSE	Q07230	mus musculu
	422	7	1.0	618	1	CHEA_LISIN	Q92dw2	listeria in
	423	7	1.0	618	1	CHEA_LISMO	Q48768	listeria mo
	424	7	1.0	622	1	FACG_HUMAN	O15287	homo sapien
	425	7	1.0	624	1	DSK1_CYLFU	Q39493	cylindrothe
c	426	7	1.0	625	1	CIKG_RAT	Q63734	rattus norv
c	427	7	1.0	627	1	K2C1_MOUSE	P04104	mus musculu
	428	7	1.0	635	1	VP40_HSV11	P10210	herpes simp
c	429	7	1.0	636	1	CTR4_HUMAN	O43246	homo sapien
	430	7	1.0	637	1	FTSH_BACSU	P37476	bacillus su
	431	7	1.0	643	1	CTK2_XENLA	P79955	xenopus lae
c	432	7	1.0	645	1	PDA4_HUMAN	P13667	homo sapien
	433	7	1.0	649	1	VATI_METMA	O59659	methanosarc
c	434	7	1.0	652	1	CD93_HUMAN	Q9npy3	homo sapien
	435	7	1.0	652	1	SSP1_SCHPO	P50526	schizosacch
c	436	7	1.0	653	1	HGFA_MOUSE	Q9r098	mus musculu
c	437	7	1.0	655	1	HGFA_HUMAN	Q04756	homo sapien
c	438	7	1.0	658	1	CNT1_RABIT	Q9mzt2	oryctolagus
	439	7	1.0	664	1	MS16_YEAST	P15424	saccharomyc
	440	7	1.0	665	1	BYR4_SCHPO	Q10951	schizosacch
c	441	7	1.0	671	1	HMOC_DROME	P22810	drosophila
	442	7	1.0	671	1	KRP2_RAT	Q62909	rattus norv
	443	7	1.0	671	1	NCPR_MUSDO	Q07994	musca domes
	444	7	1.0	671	1	PEXH_YARLI	P87200	yarrowia li
	445	7	1.0	672	1	OSM3_CAEEL	P46873	caenorhabdi
	446	7	1.0	673	1	KNS2_HUMAN	Q9bw19	homo sapien
c	447	7	1.0	675	1	HS7M_PEA	P37900	pisum sativ
	448	7	1.0	679	1	KIF2_HUMAN	O00139	homo sapien
	449	7	1.0	682	1	DYI2_CHLRE	Q39578	chlamydomon
	450	7	1.0	682	1	KIF2_XENLA	Q91637	xenopus lae
	451	7	1.0	683	1	VTER_HSVSA	Q01020	herpesvirus
c	452	7	1.0	684	1	EP84_HCMVA	P17151	human cytom
c	453	7	1.0	686	1	PLB3_YEAST	Q08108	saccharomyc
c	454	7	1.0	695	1	CIQ4_HUMAN	P56696	homo sapien
c	455	7	1.0	697	1	BYN_DROME	P55965	drosophila
	456	7	1.0	700	1	NCD_DROME	P20480	drosophila
c	457	7	1.0	708	1	GBF_DICDI	P36417	dictyosteli
c	458	7	1.0	709	1	PNP_PHOLU	P41121	photorhabdu
	459	7	1.0	716	1	KIF2_MOUSE	P28740	mus musculu
	460	7	1.0	718	1	MCAK_CRIGR	P70096	cricetulus
	461	7	1.0	722	1	SYGB_SYNY3	Q55690	synechocyst
	462	7	1.0	725	1	MCAK_HUMAN	Q99661	homo sapien
	463	7	1.0	725	1	VR2A_BPT4	P03690	bacteriopha
	464	7	1.0	729	1	KAR3_YEAST	P17119	saccharomyc
	465	7	1.0	730	1	KCM1_XENLA	Q91636	xenopus lae
c	466	7	1.0	732	1	TRI6_ECOLI	Q00191	escherichia
c	467	7	1.0	737	1	SKN1_CANAL	P87024	candida alb
	468	7	1.0	738	1	YAS9_SCHPO	Q10145	schizosacch
	469	7	1.0	740	1	DDX1_HUMAN	Q92499	homo sapien
	470	7	1.0	745	1	KATB_ARATH	P46864	arabidopsis
c	471	7	1.0	754	1	ECE1_BOVIN	P42891	bos taurus
c	472	7	1.0	754	1	ECE1_CAVPO	P97739	cavia porce
	473	7	1.0	754	1	KATC_ARATH	P46875	arabidopsis
c	474	7	1.0	759	1	CIQ2_MOUSE	Q9z351	mus musculu
c	475	7	1.0	762	1	ECE1_RAT	P42893	rattus norv
c	476	7	1.0	763	1	DYRA_HUMAN	Q13627	homo sapien
c	477	7	1.0	763	1	DYRA_MOUSE	Q61214	mus musculu
c	478	7	1.0	763	1	DYRA_RAT	Q63470	rattus norv
c	479	7	1.0	767	1	ACES_CHICK	P36196	gallus gall
c	480	7	1.0	770	1	ECE1_HUMAN	P42892	homo sapien
	481	7	1.0	770	1	KLPA_EMENI	P28739	emericella
	482	7	1.0	793	1	KATA_ARATH	Q07970	arabidopsis
c	483	7	1.0	793	1	KF3C_HUMAN	O14782	homo sapien
c	484	7	1.0	797	1	AF32_HUMAN	Q9y4w6	homo sapien
c	485	7	1.0	802	1	NAB3_YEAST	P38996	saccharomyc
c	486	7	1.0	814	1	SEFC_SALEN	P33388	salmonella
	487	7	1.0	819	1	PPSA_PYRAB	Q9v2h7	pyrococcus
	488	7	1.0	821	1	PPSA_PYRHO	O57830	pyrococcus
c	489	7	1.0	822	1	FER_HUMAN	P16591	homo sapien
	490	7	1.0	823	1	PARC_STRPN	P72525	streptococc
	491	7	1.0	832	1	KLP1_SCHPO	Q92376	schizosacch
	492	7	1.0	836	1	UME6_YEAST	P39001	saccharomyc
c	493	7	1.0	837	1	XYNZ_CLOTM	P10478	clostridium
c	494	7	1.0	842	1	SYA_CAMJE	Q9pi05	campylobact
c	495	7	1.0	844	1	DNL4_HUMAN	P49917	homo sapien
c	496	7	1.0	852	1	CIQ2_RAT	O88943	rattus norv
c	497	7	1.0	854	1	CNRB_HUMAN	P35913	homo sapien
c	498	7	1.0	854	1	VG12_BPPH2	P20345	bacteriopha
c	499	7	1.0	872	1	CIQ2_HUMAN	O43526	homo sapien
	500	7	1.0	876	1	PRO2_YEAST	P20095	saccharomyc
c	501	7	1.0	878	1	CIQ5_MOUSE	Q9jk45	mus musculu
	502	7	1.0	878	1	FIMD_ECOLI	P30130	escherichia
	503	7	1.0	884	1	SECA_OLILU	Q32743	olisthodisc
c	504	7	1.0	887	1	SM6B_RAT	O70141	rattus norv
c	505	7	1.0	890	1	ATS8_HUMAN	Q9up79	homo sapien
c	506	7	1.0	897	1	CIQ5_HUMAN	Q9nr82	homo sapien
	507	7	1.0	900	1	KDPD_CLOAB	P94608	clostridium
c	508	7	1.0	907	1	GACS_PSESY	P48027	pseudomonas
c	509	7	1.0	928	1	VGLB_MCMVS	P27171	murine cyto
	510	7	1.0	935	1	KINH_SYNRA	O43093	syncephalas
c	511	7	1.0	953	1	YA42_HUMAN	Q9upv9	homo sapien
	512	7	1.0	955	1	KINL_LEICH	P46865	leishmania
c	513	7	1.0	958	1	AMYG_DEBOC	P22861	debaryomyce
c	514	7	1.0	966	1	FIB1_PETMA	P02674	petromyzon
	515	7	1.0	970	1	T3RE_BPP1	P08764	bacteriopha
c	516	7	1.0	971	1	YQIA_CAEEL	Q09281	caenorhabdi
c	517	7	1.0	974	1	RHG6_HUMAN	O43182	homo sapien
c	518	7	1.0	981	1	GLNE_HAEIN	P44419	haemophilus
c	519	7	1.0	982	1	P115_MYCGE	P47540	mycoplasma
c	520	7	1.0	986	1	RHG6_MOUSE	O54834	mus musculu
c	521	7	1.0	992	1	SNXJ_HUMAN	Q92543	homo sapien
c	522	7	1.0	994	1	TNP5_ECOLI	P08504	escherichia
c	523	7	1.0	1004	1	PTPX_RAT	Q63475	rattus norv
	524	7	1.0	1004	1	SAL2_MOUSE	Q9qx96	mus musculu
c	525	7	1.0	1011	1	MY1A_DROME	Q23978	drosophila
	526	7	1.0	1011	1	POLG_FMDVT	P15072	foot-and-mo
	527	7	1.0	1013	1	M2B1_MOUSE	O09159	mus musculu
c	528	7	1.0	1013	1	PMPG_CHLTR	O84879	chlamydia t
	529	7	1.0	1017	1	ACAD_ARATH	Q9lik7	arabidopsis

	530	7	1.0	1020	1	V120_HSVEB	P28956	equine herp
	531	7	1.0	1023	1	GLT_DROME	P33438	drosophila
	532	7	1.0	1024	1	GCP5_HUMAN	Q96rt8	homo sapien
	533	7	1.0	1044	1	BUB1_SCHPO	O94751	schizosacch
	534	7	1.0	1063	1	ELF1_DROME	P13002	drosophila
c	535	7	1.0	1064	1	ISK5_HUMAN	Q9nq38	homo sapien
c	536	7	1.0	1078	1	GYRB_SYNY3	P77966	synechocyst
c	537	7	1.0	1086	1	DPOD_SCHPO	P30316	schizosacch
c	538	7	1.0	1093	1	AF17_HUMAN	P55198	homo sapien
	539	7	1.0	1129	1	EG27_CAEEL	Q09228	caenorhabdi
	540	7	1.0	1139	1	INA1_CAEEL	Q03600	caenorhabdi
c	541	7	1.0	1143	1	RPOB_PORPU	P51252	porphyra pu
	542	7	1.0	1146	1	CCAS_RAT	Q02485	rattus norv
c	543	7	1.0	1164	1	TSC1_HUMAN	Q92574	homo sapien
c	544	7	1.0	1216	1	RPOB_TROWH	Q93gf2	tropheryma
	545	7	1.0	1225	1	KF4A_CHICK	Q90640	gallus gall
	546	7	1.0	1226	1	KF4A_XENLA	Q91784	xenopus lae
	547	7	1.0	1231	1	KF4A_MOUSE	P33174	mus musculu
	548	7	1.0	1232	1	KF4A_HUMAN	O95239	homo sapien
c	549	7	1.0	1234	1	PIP3_MOUSE	P51432	mus musculu
c	550	7	1.0	1235	1	DNBI_HCMVA	P17147	human cytom
	551	7	1.0	1243	1	DPOA_ORYSA	O48653	oryza sativ
	552	7	1.0	1262	1	TPP2_MOUSE	Q64514	mus musculu
	553	7	1.0	1267	1	Y211_HUMAN	Q92610	homo sapien
c	554	7	1.0	1282	1	DOME_DROME	Q9vwe0	drosophila
c	555	7	1.0	1305	1	CYA9_XENLA	P98999	xenopus lae
c	556	7	1.0	1306	1	ACE_HUMAN	P12821	homo sapien
c	557	7	1.0	1319	1	MN1_HUMAN	Q10571	homo sapien
c	558	7	1.0	1324	1	SYJ1_BOVIN	O18964	bos taurus
c	559	7	1.0	1375	1	GTFC_STRMU	P13470	streptococc
c	560	7	1.0	1386	1	RPOD_MARPO	P06274	marchantia
c	561	7	1.0	1400	1	RON_HUMAN	Q04912	homo sapien
	562	7	1.0	1403	1	PRO_DROME	P29617	drosophila
c	563	7	1.0	1490	1	CRK7_HUMAN	Q9nyv4	homo sapien
c	564	7	1.0	1515	1	YCFI_YEAST	P39109	saccharomyc
	565	7	1.0	1518	1	KKK1_YEAST	P34244	saccharomyc
c	566	7	1.0	1558	1	YK83_YEAST	P36028	saccharomyc
	567	7	1.0	1559	1	BPT1_YEAST	P14772	saccharomyc
c	568	7	1.0	1575	1	SYJ1_HUMAN	O43426	homo sapien
c	569	7	1.0	1620	1	ALK_HUMAN	Q9um73	homo sapien
c	570	7	1.0	1621	1	ALK_MOUSE	P97793	mus musculu
	571	7	1.0	1647	1	SN24_HUMAN	P51532	homo sapien
c	572	7	1.0	1663	1	CO3_HUMAN	P01024	homo sapien
	573	7	1.0	1709	1	CHD1_HUMAN	O14646	homo sapien
	574	7	1.0	1711	1	CHD1_MOUSE	P40201	mus musculu
	575	7	1.0	1739	1	CHD2_HUMAN	O14647	homo sapien
c	576	7	1.0	1785	1	BIG2_HUMAN	Q9y6d5	homo sapien
	577	7	1.0	1852	1	CCAS_CYPCA	P22316	cyprinus ca
	578	7	1.0	1873	1	CCAS_HUMAN	Q13698	homo sapien
	579	7	1.0	1873	1	CCAS_RABIT	P07293	oryctolagus
c	580	7	1.0	1912	1	PTPD_HUMAN	P23468	homo sapien
c	581	7	1.0	1938	1	MYS_AEQIR	P24733	aequipecten
c	582	7	1.0	2038	1	FSH_DROME	P13709	drosophila
	583	7	1.0	2116	1	MYS2_DICDI	P08799	dictyosteli
c	584	7	1.0	2167	1	BEM2_YEAST	P39960	saccharomyc
	585	7	1.0	2193	1	POLG_CXA16	Q65900	c genome po
c	586	7	1.0	2209	1	Y166_HUMAN	P50748	homo sapien
	587	7	1.0	2222	1	DPOE_YEAST	P21951	saccharomyc
	588	7	1.0	2332	1	POLG_FMDVA	P03308	f genome po
	589	7	1.0	2332	1	POLG_FMDVO	P03305	f genome po
	590	7	1.0	2333	1	POLG_FMDV1	P03306	f genome po
	591	7	1.0	2336	1	POLG_FMDVZ	P49303	f genome po
c	592	7	1.0	2368	1	ESR1_YEAST	P38111	saccharomyc
c	593	7	1.0	2415	1	SPCA_MOUSE	P08032	mus musculu
c	594	7	1.0	2517	1	NCR2_HUMAN	Q9y618	h nuclear r
c	595	7	1.0	3005	1	ZFH2_DROME	P28167	drosophila
c	596	7	1.0	4367	1	DYHC_NEUCR	P45443	neurospora
	597	7	1.0	4427	1	PKSL_BACSU	Q05470	bacillus su
c	598	7	1.0	4660	1	LRP2_RAT	P98158	rattus norv
c	599	7	1.0	5035	1	RYR1_PIG	P16960	sus scrofa
c	600	7	1.0	5037	1	RYR1_RABIT	P11716	oryctolagus
c	601	7	1.0	5038	1	RYR1_HUMAN	P21817	homo sapien
	602	6	0.8	15	1	CXA1_CONGE	P01519	conus geogr

c	603	6	0.9	31	1	LPL_BUCRP	Q53017	buchnera ap
	604	6	0.8	32	1	CAL_BOVIN	P01260	bos taurus
	605	6	0.8	32	1	CAL_PIG	P01259	sus scrofa
c	606	6	0.9	32	1	MDH_NITAL	P10887	nitzschia a
	607	6	0.8	32	1	PRT4_SCYCA	P30259	scyliorhinu
	608	6	0.8	37	1	PETG_CUSRE	P30398	cuscuta ref
c	609	6	0.9	40	1	VIT_MELGA	P56531	meleagris g
c	610	6	0.9	42	1	PSBK_CHLVU	P56348	chlorella v
c	611	6	0.9	43	1	MUTI_ENTMU	P80925	enterococcu
c	612	6	0.9	44	1	PSBK_ODOSI	P49513	odontella s
c	613	6	0.9	45	1	PSBK_EUGVI	Q9ms55	euglena vir
c	614	6	0.9	46	1	PSBK_CHLRE	P18263	chlamydomon
c	615	6	0.9	48	1	PSBK_EUGSA	Q9ms61	euglena san
c	616	6	0.9	49	1	R331_STRPN	Q97r59	streptococc
	617	6	0.8	50	1	HSP1_HUMAN	P04553	homo sapien
	618	6	0.8	50	1	HSP1_HYLLA	P35306	hylobates l
	619	6	0.8	53	1	RL2A_PIG	Q29333	sus scrofa
c	620	6	0.9	54	1	B2MG_MESAU	P30442	mesocricetu
	621	6	0.8	54	1	GALA_CANFA	P33710	canis famil
	622	6	0.8	56	1	HSP_COTJA	P14402	coturnix co
	623	6	0.8	57	1	HSP1_DIDMA	P35305	didelphis m
c	624	6	0.9	57	1	LHA_CHLAU	P07503	chloroflexu
c	625	6	0.9	57	1	RK32_CHLVU	P56357	chlorella v
c	626	6	0.9	58	1	RL32_NEIMA	Q9jw52	neisseria m
c	627	6	0.9	58	1	RL32_NEIMB	Q9jxs0	neisseria m
	628	6	0.8	59	1	HSP1_PERCO	Q9glq5	peradorcas
	629	6	0.8	60	1	HSP1_AEPRU	Q9glp9	aepyprymnus
	630	6	0.8	60	1	HSP1_CAEFU	P42131	caenolestes
	631	6	0.8	60	1	HSP1_DENGO	Q9gjq1	dendrolagus
	632	6	0.8	60	1	HSP1_MACAG	P42137	macropus ag
	633	6	0.8	60	1	HSP1_MACGI	P42139	macropus gi
	634	6	0.8	60	1	HSP1_MACPA	Q9glq9	macropus pa
	635	6	0.8	60	1	HSP1_ONCFR	Q9glq3	onychogalea
	636	6	0.8	60	1	HSP1_POTLO	Q9glp8	potorous lo
c	637	6	0.9	61	1	DNBI_BFDV	P13893	budgerigar
	638	6	0.8	61	1	HSP1_BETPE	Q9glq0	bettongia p
	639	6	0.8	61	1	HSP1_MACEU	P42138	macropus eu
	640	6	0.8	61	1	HSP1_MACRG	P42141	macropus ru
	641	6	0.8	61	1	HSP1_PETXA	Q9glq7	petrogale x
	642	6	0.8	61	1	HSP1_THYST	Q9glq8	thylogale s
	643	6	0.8	61	1	HSP1_TRIVU	P42152	trichosurus
	644	6	0.8	61	1	HSP_CHICK	P15340	gallus gall
c	645	6	0.9	61	1	PSBK_LOTJA	Q9bbs2	lotus japon
	646	6	0.8	61	1	RL21_PIG	P49666	sus scrofa
c	647	6	0.9	62	1	YPPF_BACSU	P50834	bacillus su
	648	6	0.8	63	1	HSP1_HYPMS	Q9glq1	hypsiprymno
	649	6	0.8	63	1	MIP_BOTAS	P81077	bothrops as
	650	6	0.8	64	1	HSP1_LAGHI	Q9glq2	lagorcheste
	651	6	0.8	64	1	RL35_STRCO	O88059	streptomyce
	652	6	0.8	65	1	MYHB_PIG	P81271	sus scrofa
c	653	6	0.9	65	1	PHX2_MOUSE	P15972	mus musculu
c	654	6	0.9	65	1	TXM7_DENAN	P80970	dendroaspis
	655	6	0.8	66	1	YPJJ_ECOLI	P58033	escherichia
	656	6	0.8	67	1	CSPA_MICLU	O30875	micrococcus
	657	6	0.8	68	1	HSP1_PSECU	P42145	pseudochiro
	658	6	0.8	68	1	PLE2_PSEAM	Q9prj9	pseudopleur
	659	6	0.8	70	1	L2MU_ADE40	Q64858	human adeno
c	660	6	0.9	70	1	PYRH_THETH	P43891	thermus the
c	661	6	0.9	70	1	YIIE_ECOLI	P32149	escherichia
	662	6	0.8	71	1	DNBI_POVJC	P03086	polyomaviru
	663	6	0.8	71	1	LEA2_CAVPO	Q91x13	cavia porce
c	664	6	0.9	71	1	YDAQ_ECOLI	P76057	escherichia
	665	6	0.8	72	1	VP8_TCV	P17461	turnip crin
c	666	6	0.9	74	1	SRP_SOYBN	Q07502	glycine max
	667	6	0.8	74	1	UL11_HSVEB	P28982	equine herp
	668	6	0.8	75	1	YF45_MYCTU	Q10781	mycobacteri
c	669	6	0.9	76	1	CXO3_CONTE	Q9xzl0	conus texti
	670	6	0.8	77	1	DIIR_ECOLI	P21320	escherichia
c	671	6	0.9	77	1	ENV_SMSAV	P03384	simian sarc
	672	6	0.8	77	1	LEA2_BOVIN	Q95jc3	bos taurus
	673	6	0.8	77	1	PRT2_SEPOF	P80002	sepia offic
	674	6	0.8	77	1	YF78_ARCFU	O28694	archaeoglob
c	675	6	0.9	77	1	YF78_ARCFU	O28694	archaeoglob

676	6	0.8	78	1	PRT1_SEPOF	P80001 sepia offic
677	6	0.8	78	1	VE5_HPV41	P27554 human papil
678	6	0.8	79	1	CSMA_CHLAU	P09928 chloroflexu
679	6	0.8	79	1	EFG_BACST	P23081 bacillus st
c 680	6	0.9	79	1	OAG3_SALTI	Q8z9m5 salmonella
c 681	6	0.9	79	1	OAG3_SALTY	P58651 salmonella
682	6	0.8	79	1	PRL3_HUMAN	P02814 homo sapien
c 683	6	0.9	80	1	OAG1_SALTI	Q8z3e4 salmonella
c 684	6	0.9	80	1	YDWA_SCHPO	O13915 schizosacch
c 685	6	0.9	80	1	YDWC_SCHPO	O13916 schizosacch
c 686	6	0.9	80	1	YHY5_WOLSU	P31877 wolinella s
c 687	6	0.9	81	1	LUXC_VIBFI	P12748 vibrio fisc
c 688	6	0.9	81	1	OAG2_SALTY	P58650 salmonella
689	6	0.8	81	1	POC3_SYRVU	P58171 syringa vul
690	6	0.8	81	1	YQGY_BACSU	P54502 bacillus su
691	6	0.8	82	1	CYTA_HELAN	Q10992 helianthus
692	6	0.8	82	1	P8_HUMAN	O60356 homo sapien
693	6	0.8	83	1	COXG_SCHPO	O94581 schizosacch
c 694	6	0.9	83	1	HEPC_MOUSE	Q9eq21 mus musculu
695	6	0.8	83	1	V187_BPT3	P10302 bacteriopha
c 696	6	0.9	83	1	VG03_BPMD2	O64199 mycobacteri
c 697	6	0.9	84	1	BCE1_HUMAN	O60756 homo sapien
c 698	6	0.9	84	1	COAB_BPI22	P15416 bacteriopha
c 699	6	0.9	84	1	HEPC_HUMAN	P81172 homo sapien
c 700	6	0.9	84	1	HEPC_RAT	Q99mh3 rattus norv
701	6	0.8	84	1	POC3_OLEEU	O81092 olea europa
702	6	0.8	84	1	RL24_SULAC	O05633 sulfolobus
703	6	0.8	84	1	UNGI_BPPB2	P14739 bacteriopha
704	6	0.8	84	1	YEXA_BACSU	P12049 bacillus su
c 705	6	0.9	84	1	YF7A_PSEAE	P58040 pseudomonas
c 706	6	0.9	85	1	FCEG_BOVIN	Q9bdr7 bos taurus
c 707	6	0.9	85	1	HIS3_SULSO	O33778 sulfolobus
c 708	6	0.9	85	1	MTRG_METTH	O27225 methanobact
c 709	6	0.9	85	1	MTRG_METTM	Q50774 methanobact
c 710	6	0.9	86	1	FCEG_CAVPO	Q07249 cavia porce
c 711	6	0.9	86	1	FCEG_HUMAN	P30273 homo sapien
c 712	6	0.9	86	1	FCEG_PIG	Q9xsz6 sus scrofa
713	6	0.8	86	1	PAK1_SYNY3	P72712 synechocyst
c 714	6	0.9	86	1	RL23_METJA	P54016 methanococc
c 715	6	0.9	87	1	IM08_YEAST	P57744 saccharomyc
c 716	6	0.9	90	1	RK16_OENAM	P42355 oenothera a
c 717	6	0.9	90	1	VG09_BPPF1	P25135 bacteriopha
c 718	6	0.9	91	1	DBH_STRTR	P96045 streptococc
c 719	6	0.9	91	1	V179_FOWPV	Q9j554 fowlpox vir
720	6	0.8	92	1	Y432_METJA	Q57874 methanococc
c 721	6	0.9	92	1	YE21_HAEIN	P44192 haemophilus
722	6	0.8	93	1	RT19_MARPO	P26874 marchantia
c 723	6	0.9	93	1	UGR1_HUMAN	Q96pl1 homo sapien
c 724	6	0.9	93	1	Y435_METJA	Q57877 methanococc
c 725	6	0.9	94	1	YIF8_YEAST	P40521 saccharomyc
c 726	6	0.9	95	1	YCNE_BACSU	P94425 bacillus su
727	6	0.8	95	1	YPSS_RHILP	P10497 rhizobium l
c 728	6	0.9	96	1	NU6M_ALBTU	Q08084 albinaria t
c 729	6	0.9	97	1	7E_DICDI	P15649 dictyosteli
730	6	0.8	97	1	CYTI_VIGUN	Q06445 vigna ungui
731	6	0.8	97	1	DEF4_HUMAN	P12838 homo sapien
c 732	6	0.9	97	1	IM8A_RAT	Q9wva1 rattus norv
733	6	0.8	97	1	VE7_HPV23	P50781 human papil
734	6	0.8	97	1	VPR_SIVMK	P05957 simian immu
735	6	0.8	97	1	YE20_HAEIN	P44191 haemophilus
736	6	0.8	97	1	YXYB_CALSA	P23554 caldocellum
737	6	0.8	98	1	KRFT_LARNO	P02451 larus novae
c 738	6	0.9	98	1	NULM_LUMTE	Q34948 lumbricus t
c 739	6	0.9	98	1	SNC5_HUMAN	O75971 homo sapien
c 740	6	0.9	98	1	V55_BPT3	P20319 bacteriopha
c 741	6	0.9	99	1	B2MG_RABIT	P01885 oryctolagus
c 742	6	0.9	99	1	SR19_PYRAB	Q9v1j9 pyrococcus
c 743	6	0.9	99	1	SR19_PYRHO	O74102 pyrococcus
c 744	6	0.9	99	1	SZ14_HUMAN	O95715 homo sapien
c 745	6	0.9	99	1	SZ14_MOUSE	Q9wuq5 mus musculu
c 746	6	0.9	99	1	YHCO_BACSU	P54599 bacillus su
c 747	6	0.9	100	1	BOP1_BOVIN	P83107 bos taurus
c 748	6	0.9	100	1	MI2B_RAT	Q10747 rattus norv
c 749	6	0.9	100	1	RL23_ACTAC	P55839 actinobacil
c 750	6	0.9	100	1	TXOA_ATRIL	Q9bjv8 atrax sp. i
c 751	6	0.9	100	1	TXOB_ATRIL	Q9bjv7 atrax sp. i
752	6	0.8	100	1	YGFE_HAEIN	P44062 haemophilus
753	6	0.8	100	1	YMTS_CLOTT	Q05500 clostridium
c 754	6	0.9	101	1	GRO_CRIGR	P09340 cricetulus
755	6	0.8	101	1	H1L6_ENSMI	P27204 ensis minor
c 756	6	0.9	101	1	RS16_UREPA	Q9pps1 ureaplasma
757	6	0.8	101	1	VPR_SIVM1	P05958 simian immu
758	6	0.8	102	1	CYTI_ORYSA	P09229 oryza sativ
759	6	0.8	102	1	YBJ6_YEAST	P38171 saccharomyc
c 760	6	0.9	103	1	ATPN_HUMAN	O75964 homo sapien
c 761	6	0.9	103	1	MGP_RAT	P08494 rattus norv
c 762	6	0.9	103	1	SMS2_RANRI	P87385 rana ridibu
c 763	6	0.9	103	1	YGU3_YEAST	P53097 saccharomyc
c 764	6	0.9	104	1	GRO2_RABIT	P47854 oryctolagus
c 765	6	0.9	104	1	GRO_CAVPO	O55235 cavia porce
766	6	0.8	104	1	HSP2_CALJA	Q28337 callithrix
767	6	0.8	104	1	HSP2_RAT	P11248 rattus norv
c 768	6	0.9	104	1	MGP_MOUSE	P19788 mus musculu
769	6	0.8	104	1	VPR_HV2ST	P20884 human immun
c 770	6	0.9	105	1	CORT_HUMAN	O00230 homo sapien
c 771	6	0.9	105	1	RLA2_LEIBR	O44010 leishmania
c 772	6	0.9	105	1	VMEM_FXMV	P22170 foxtail mos
773	6	0.8	105	1	VPR_HV2SB	P12455 human immun
774	6	0.8	105	1	YB59_YEAST	P38311 saccharomyc
775	6	0.8	105	1	YH06_METTH	O27741 methanobact
776	6	0.8	106	1	YCIH_HAEIN	P45116 haemophilus
777	6	0.8	107	1	CYT2_ORYSA	P20907 oryza sativ
778	6	0.8	107	1	ELBA_ECOLI	P75987 escherichia
c 779	6	0.9	107	1	GRO_HUMAN	P09341 homo sapien
780	6	0.8	107	1	HSP2_MOUSE	P07978 mus musculu
c 781	6	0.9	107	1	MI2A_HUMAN	P19875 homo sapien
c 782	6	0.9	107	1	MI2B_HUMAN	P19876 homo sapien
c 783	6	0.9	107	1	PRK2_RAT	Q8r413 rattus norv
c 784	6	0.9	107	1	Y240_RICPR	Q9zdt5 rickettsia
c 785	6	0.9	107	1	Y324_RICCN	Q92iu6 rickettsia
c 786	6	0.9	108	1	SVS4_MOUSE	P18419 mus musculu
787	6	0.8	108	1	Y14A_BPT4	P32279 bacteriopha
788	6	0.8	108	1	Y187_BUCAI	P57284 buchnera ap
789	6	0.8	108	1	YCGL_ECOLI	P76003 escherichia
790	6	0.8	108	1	YN1I_CAEEL	P91375 caenorhabdi
c 791	6	0.9	109	1	ZM33_MAIZE	O82106 zea mays (m
c 792	6	0.9	110	1	CU15_HUMAN	P58622 homo sapien
793	6	0.8	110	1	YE42_THEMA	Q9x1f5 thermotoga
794	6	0.8	111	1	YRZC_BACSU	O34527 bacillus su
c 795	6	0.9	112	1	CORT_RAT	Q62949 rattus norv
c 796	6	0.9	112	1	OL7F_MOUSE	Q60887 mus musculu
797	6	0.8	112	1	SVS4_RAT	P02783 rattus norv
c 798	6	0.9	112	1	SVS4_RAT	P02783 rattus norv
c 799	6	0.9	112	1	TTF1_CAVPO	P97273 cavia porce
c 800	6	0.9	112	1	YBV0_YEAST	P38259 saccharomyc
c 801	6	0.9	113	1	KV2D_HUMAN	P01617 homo sapien
802	6	0.8	114	1	IR03_HCMVA	P17142 human cytom
c 803	6	0.9	114	1	NPFF_MOUSE	Q9wva8 mus musculu
c 804	6	0.9	114	1	NPFF_RAT	Q9wva9 rattus norv
c 805	6	0.9	114	1	SZ05_HUMAN	P42830 homo sapien
c 806	6	0.9	114	1	SZ06_HUMAN	P80162 homo sapien
c 807	6	0.9	115	1	ATKA_BACST	P94456 b potassium
808	6	0.8	115	1	HIS3_MYCLE	Q9x7c3 mycobacteri
809	6	0.8	115	1	HIS3_MYCTU	O53909 mycobacteri
c 810	6	0.9	115	1	RA2A_ARATH	Q9slf7 arabidopsis
c 811	6	0.9	115	1	RA2B_ARATH	P51407 arabidopsis
c 812	6	0.9	115	1	SMS1_PROAN	Q9w7f0 protopterus
c 813	6	0.9	115	1	Y440_ARCFU	O29809 archaeoglob
c 814	6	0.9	115	1	YN43_YEAST	P53827 saccharomyc
815	6	0.8	115	1	YSCI_YEREN	Q01250 yersinia en
816	6	0.8	115	1	YSCI_YERPE	Q00933 yersinia pe
817	6	0.8	116	1	GLB_PARCA	P15160 paramecium
818	6	0.8	116	1	NU3M_ASTPE	P11991 asterina pe
c 819	6	0.9	116	1	NU3M_STRCA	O79102 struthio ca
820	6	0.8	116	1	Y960_HAEIN	P44084 haemophilus
c 821	6	0.9	116	1	YBQ1_YEAST	P38233 saccharomyc

c	822	6	0.9	117	1	GP49_BPSP1	O48403 bacteriopha
c	823	6	0.9	117	1	KV2E_HUMAN	P06309 homo sapien
c	824	6	0.9	117	1	RNPA_LACLA	Q9cj73 lactococcus
c	825	6	0.9	117	1	UREF_BACPA	Q45345 bacillus pa
c	826	6	0.9	117	1	WAP_CAMDR	P09837 camelus dro
	827	6	0.8	118	1	RR19_ASTLO	P34772 astasia lon
c	828	6	0.9	118	1	YAI9_SCHPO	Q09896 schizosacch
c	829	6	0.9	119	1	HIS3_SULTO	Q970y6 sulfolobus
c	830	6	0.9	119	1	RLA3_MAIZE	O24413 zea mays (m
	831	6	0.8	119	1	RNL4_BOVIN	P15467 bos taurus
c	832	6	0.9	119	1	RNPA_BACHD	Q9rca4 bacillus ha
c	833	6	0.9	119	1	SZ07_PIG	P43030 sus scrofa
	834	6	0.8	119	1	WNT4_PLEJO	P28135 plethodon j
c	835	6	0.9	119	1	YJ98_YEAST	P47162 saccharomyc
c	836	6	0.9	120	1	CHH1_PENJP	O15980 penaeus jap
c	837	6	0.9	120	1	CHH2_PENJP	Q9u5d2 penaeus jap
c	838	6	0.9	120	1	GLHA_CALJA	P51499 c glycoprot
c	839	6	0.9	120	1	SY27_MOUSE	Q9z1x0 mus musculu
c	840	6	0.9	120	1	TRI7_HUMAN	Q15651 homo sapien
	841	6	0.8	120	1	Y4NI_RHISN	P55581 rhizobium s
c	842	6	0.9	121	1	AMEL_ORNAN	O97646 ornithorhyn
c	843	6	0.9	121	1	AMEL_TACAC	O97647 tachyglossu
	844	6	0.8	121	1	RL20_CHLPN	Q9z6r7 chlamydia p
c	845	6	0.9	121	1	SECR_HUMAN	P09683 homo sapien
	846	6	0.8	121	1	YGI8_YEAST	P53151 saccharomyc
	847	6	0.8	122	1	RK14_PINTH	P41633 pinus thunb
c	848	6	0.9	123	1	C59A_MOUSE	O55186 mus musculu
c	849	6	0.9	123	1	RNPA_STRPN	Q97ni5 streptococc
c	850	6	0.9	123	1	UR2_MOUSE	Q9qzq3 mus musculu
c	851	6	0.9	123	1	UR2_RAT	Q9qzq4 rattus norv
c	852	6	0.9	123	1	WNT7_STRPU	P28098 strongyloce
c	853	6	0.9	123	1	YPPE_BACSU	P50833 bacillus su
	854	6	0.8	124	1	R22B_ARATH	Q9m9w1 arabidopsis
	855	6	0.8	124	1	R22C_ARATH	Q9fe58 arabidopsis
c	856	6	0.9	124	1	RNP_GIRCA	P00662 giraffa cam
	857	6	0.8	124	1	WNT3_EVATR	P28090 evasterias
c	858	6	0.9	125	1	ALR_MOUSE	P56213 mus musculu
c	859	6	0.9	125	1	ALR_RAT	Q63042 rattus norv
c	860	6	0.9	125	1	RS13_RICCN	Q92gy8 rickettsia
	861	6	0.8	125	1	RS13_RICPR	Q9zcs7 rickettsia
c	862	6	0.9	125	1	RS13_RICPR	Q9zcs7 rickettsia
	863	6	0.8	125	1	RS6_CAMJE	Q9zah3 campylobact
	864	6	0.8	125	1	YAI7_CAEEL	Q20507 caenorhabdi
	865	6	0.8	126	1	RBFA_TREPA	O83860 treponema p
c	866	6	0.9	127	1	COAT_BPPP7	P03630 bacteriopha
c	867	6	0.9	127	1	CRCB_CAUCR	Q9a6v2 caulobacter
c	868	6	0.9	127	1	WAP_RABIT	P09412 oryctolagus
c	869	6	0.9	128	1	IF5A_ARCFU	O29612 archaeoglob
c	870	6	0.9	128	1	PRK2_MOUSE	Q9qxu7 mus musculu
c	871	6	0.9	128	1	YS98_MYCTU	Q10819 mycobacteri
	872	6	0.8	129	1	AAK2_PIG	Q28948 sus scrofa
c	873	6	0.9	129	1	C59B_MOUSE	P58019 mus musculu
	874	6	0.8	129	1	LYC2_CAPHI	P37714 capra hircu
	875	6	0.8	129	1	LYC_SHEEP	P17607 ovis aries
c	876	6	0.9	129	1	PRK2_HUMAN	Q9hc23 homo sapien
c	877	6	0.9	129	1	TKN1_HUMAN	P20366 homo sapien
c	878	6	0.9	129	1	YLS1_CAEEL	P34386 caenorhabdi
	879	6	0.8	130	1	CW36_YEAST	P25603 saccharomyc
	880	6	0.8	130	1	GTH1_CARAU	Q98848 carassius a
c	881	6	0.9	130	1	OREX_CANFA	Q9glf6 canis famil
c	882	6	0.9	130	1	OREX_MOUSE	O55241 mus musculu
c	883	6	0.9	130	1	OREX_RAT	O55232 rattus norv
c	884	6	0.9	130	1	TKN1_MESAU	Q60541 mesocricetu
c	885	6	0.9	130	1	TKN1_MOUSE	P41539 mus musculu
c	886	6	0.9	130	1	TKN1_RAT	P06767 rattus norv
	887	6	0.8	130	1	YHA5_YEAST	P38752 saccharomyc
c	888	6	0.9	131	1	OREX_HUMAN	O43612 homo sapien
c	889	6	0.9	131	1	OREX_PIG	O77668 sus scrofa
c	890	6	0.9	131	1	RS11_HELPJ	Q9zjt3 helicobacte
c	891	6	0.9	131	1	RS11_HELPY	P56018 helicobacte
c	892	6	0.9	131	1	SECR_PIG	P01279 sus scrofa
c	893	6	0.9	131	1	Y209_PYRHO	O57948 pyrococcus
	894	6	0.8	131	1	YAOG_SCHPO	Q10095 schizosacch

	895	6	0.8	132	1	ATPE_CYAPA	P48083 cyanophora
c	896	6	0.9	132	1	Y125_AQUAE	O66525 aquifex aeo
c	897	6	0.9	132	1	YVGA_VACCC	P20565 vaccinia vi
	898	6	0.8	133	1	FOLB_MYCTU	O06275 mycobacteri
c	899	6	0.9	133	1	NIKR_SULSO	Q9ux49 sulfolobus
c	900	6	0.9	133	1	SECR_MOUSE	Q08535 mus musculu
c	901	6	0.9	133	1	YJZ3_YEAST	P47094 saccharomyc
c	902	6	0.9	134	1	SECR_RAT	P11384 rattus norv
	903	6	0.8	134	1	STPA_ECOLI	P30017 escherichia
	904	6	0.8	134	1	YFF6_SCHPO	O14069 schizosacch
c	905	6	0.9	134	1	YFID_BACSU	P54720 bacillus su
c	906	6	0.9	135	1	ANFC_SQUAC	P41319 squalus aca
c	907	6	0.9	135	1	RK16_ARATH	P56793 arabidopsis
c	908	6	0.9	135	1	YGDB_ECOLI	P08370 escherichia
	909	6	0.8	136	1	GLB3_CHITP	P22431 chironomus
c	910	6	0.9	136	1	RK16_MAIZE	P08528 zea mays (m
c	911	6	0.9	136	1	RK16_ORYSA	P12138 oryza sativ
	912	6	0.8	137	1	FER1_ASCSU	P49671 ascaris suu
c	913	6	0.9	137	1	RS9_SYNY3	P73293 synechocyst
	914	6	0.8	138	1	GCSH_PYRAB	Q9v0g1 pyrococcus
c	915	6	0.9	138	1	LSHB_MACRU	O46483 macropus ru
	916	6	0.8	138	1	RS8_THETH	P24319 thermus the
	917	6	0.8	138	1	TETD_ECOLI	P28816 escherichia
c	918	6	0.9	139	1	UGR1_MOUSE	Q920h1 mus musculu
c	919	6	0.9	139	1	YN06_YEAST	P53842 saccharomyc
	920	6	0.8	139	1	YO72_RHIME	P42879 rhizobium m
c	921	6	0.9	140	1	ANF_CHICK	P18908 gallus gall
c	922	6	0.9	140	1	MOPB_ERWCA	P34199 erwinia car
	923	6	0.8	140	1	VE6_HPV04	Q07854 human papil
	924	6	0.8	140	1	VE6_HPV65	Q07856 human papil
	925	6	0.8	140	1	Y454_MYCPN	P75123 mycoplasma
	926	6	0.8	140	1	YA00_RHIME	Q52967 rhizobium m
c	927	6	0.9	140	1	YGK8_YEAST	P53139 saccharomyc
	928	6	0.8	140	1	YJ14_YEAST	P47111 saccharomyc
	929	6	0.8	141	1	DUT_CHVP1	O41033 paramecium
	930	6	0.8	141	1	HBA_EUDSC	P19789 eudynamys s
c	931	6	0.9	141	1	LSHB_HUMAN	P01229 homo sapien
c	932	6	0.9	141	1	LSHB_TRIVU	O46482 trichosurus
c	933	6	0.9	141	1	NIKR_METJA	Q57969 methanococc
c	934	6	0.9	142	1	CU07_MOUSE	P58500 mus musculu
c	935	6	0.9	142	1	HBA_CHEKU	P80270 chelodonich
c	936	6	0.9	142	1	IL3_CALJA	Q28334 callithrix
c	937	6	0.9	142	1	IL3_SAGOE	P51445 saguinus oe
c	938	6	0.9	142	1	MXIM_SHIFL	Q06083 shigella fl
	939	6	0.8	142	1	VE6_HPV48	Q80920 human papil
c	940	6	0.9	142	1	YECG_ECOLI	P46888 escherichia
	941	6	0.8	143	1	CAL_SHEEP	P01261 ovis aries
c	942	6	0.9	143	1	IL3_MACMU	P25140 macaca mula
	943	6	0.8	143	1	RK16_MARPO	P06383 marchantia
	944	6	0.8	143	1	Y363_AQUAE	O66688 aquifex aeo
	945	6	0.8	143	1	YTXG_BACSU	P40779 bacillus su
c	946	6	0.9	144	1	CSF2_HUMAN	P04141 homo sapien
c	947	6	0.9	144	1	IL3_BOVIN	P49875 bos taurus
	948	6	0.8	144	1	RIBI_PHOLE	Q01994 photobacter
	949	6	0.8	144	1	SOD1_HALME	O08461 halobacteri
	950	6	0.8	144	1	YJEK_BUCAP	Q44634 buchnera ap
	951	6	0.8	145	1	ANGR_MOUSE	Q64438 mus musculu
c	952	6	0.9	145	1	COPY_ENTHR	Q47839 enterococcu
c	953	6	0.9	145	1	PA2M_CAVPO	P47711 cavia porce
c	954	6	0.9	145	1	RL9_UREPA	Q9ppt9 ureaplasma
	955	6	0.8	145	1	SSRP_MYCPU	Q98qk9 mycoplasma
c	956	6	0.9	145	1	YA01_RHIME	Q52968 rhizobium m
	957	6	0.8	145	1	YG13_SYNY3	P73841 synechocyst
c	958	6	0.9	146	1	FUR2_MYCFO	O69451 mycobacteri
c	959	6	0.9	146	1	IL3_SHEEP	Q06435 ovis aries
	960	6	0.8	146	1	XKDJ_BACSU	P54330 bacillus su
	961	6	0.8	146	1	YD82_METJA	Q58777 methanococc
	962	6	0.8	146	1	YG5G_YEAST	P53321 saccharomyc
c	963	6	0.9	147	1	CST9_HUMAN	Q9h4g1 homo sapien
c	964	6	0.9	147	1	FUR_MYCTU	O07724 mycobacteri
c	965	6	0.9	147	1	LYC_COTJA	P00701 coturnix co
c	966	6	0.9	147	1	MCP_MEDSA	P42067 medicago sa
	967	6	0.8	147	1	PLIA_AGKBL	P82142 agkistrodon

```
   968      6   0.8    147  1  PLIA_TRIFL         P21755 trimeresuru
   969      6   0.8    147  1  PLIB_TRIFL         P21756 trimeresuru
   970      6   0.8    147  1  RL2A_HUMAN         P46776 homo sapien
   971      6   0.8    147  1  RL2A_MOUSE         P14115 mus musculu
   972      6   0.8    147  1  RL2A_RAT           P18445 rattus norv
   973      6   0.8    147  1  RL2A_XENLA         P47830 xenopus lae
   974      6   0.8    147  1  RNL4_HUMAN         P34096 homo sapien
   975      6   0.8    147  1  RNL4_PIG           P15468 sus scrofa
   976      6   0.8    147  1  YZ36_METJA         Q60257 methanococc
c  977      6   0.9    148  1  GVPJ_ANASP         Q8yut1 anabaena sp
c  978      6   0.9    148  1  NDKB_FLABI         P47920 flaveria bi
   979      6   0.8    148  1  NIKR_HELPJ         Q9zjp7 helicobacte
   980      6   0.8    148  1  NIKR_HELPY         O25896 helicobacte
   981      6   0.8    148  1  RL2A_DICDI         P48160 dictyosteli
   982      6   0.8    148  1  RL2A_EUPCR         P48161 euplotes cr
   983      6   0.8    148  1  RL2A_TENMO         Q27021 tenebrio mo
c  984      6   0.9    148  1  SY02_MOUSE         P10148 mus musculu
c  985      6   0.9    148  1  SY02_RAT           P14844 rattus norv
c  986      6   0.9    148  1  YCXJ_PORPU         P51350 porphyra pu
c  987      6   0.9    149  1  LRPB_BACSU         P96653 bacillus su
c  988      6   0.9    149  1  NDK1_ORYSA         Q07661 oryza sativ
   989      6   0.8    149  1  RBFA_CAUCR         Q9ac30 caulobacter
   990      6   0.8    149  1  RL2A_DROME         P41092 drosophila
   991      6   0.8    149  1  RL9_HELPJ          Q9zlw3 helicobacte
   992      6   0.8    149  1  RL9_HELPY          P56035 helicobacte
c  993      6   0.9    149  1  YFEZ_ECOLI         P76538 escherichia
c  994      6   0.9    150  1  AVR1_CHICK         O13153 gallus gall
c  995      6   0.9    150  1  AVR2_CHICK         P56732 gallus gall
c  996      6   0.9    150  1  AVR3_CHICK         P56733 gallus gall
c  997      6   0.9    150  1  AVR4_CHICK         P56734 gallus gall
c  998      6   0.9    150  1  AVR6_CHICK         P56735 gallus gall
c  999      6   0.9    150  1  AVR7_CHICK         P56736 gallus gall
c1000       6   0.9    150  1  RNK6_CERAE         O46534 cercopithec
```

ALIGNMENTS

```
RESULT 1
DDX4_HUMAN
ID   DDX4_HUMAN     STANDARD;      PRT;   724 AA.
AC   Q9NQI0; Q9NYB1; Q9NT92;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 4 (VASA homolog).
GN   DDX4 OR VASA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC   TISSUE=Testis;
RX   MEDLINE=20402578; PubMed=10920202;
RA   Castrillon D.H., Quade B.J., Wang T.Y., Quigley C., Crum C.P.;
RT   "The human VASA gene is specifically expressed in the germ cell
RT   lineage.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:9585-9590(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Rocha D., Affara N.;
RT   "Cloning and characterization of the human VASA gene.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 90-724 FROM N.A.
RC   TISSUE=Testis;
RA   Blum H., Bauersachs S., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May play a role in germ cell development.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Expressed only in ovary and testis. Expressed
CC       in migratory primordial germ cells in the region of the gonadal
CC       ridge in both sexes.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC       DDX4/VASA SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AY004154; AAF86585.1; -.
DR   EMBL; AF262962; AAF72705.1; -.
DR   EMBL; AL137462; CAB70750.1; -.
DR   HSSP; Q58083; 1HV8.
DR   Genew; HGNC:18700; DDX4.
DR   MIM; 605281; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 2.
DR   Pfam; PF00271; helicase_C; 2.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.
FT   NP_BIND     332    339       ATP (POTENTIAL).
FT   SITE        446    449       DEAD BOX.
FT   DOMAIN       58    234       GLY-RICH.
FT   CONFLICT    556    556       I -> T (IN REF. 1).
FT   CONFLICT    713    713       Q -> R (IN REF. 1).
SQ   SEQUENCE   724 AA;  79307 MW;  6D313DD98B177067 CRC64;

Alignment Scores:
Pred. No.:              0              Length:        724
Score:                  611.00         Matches:       711
Percent Similarity:     99.86%         Conservative:  0
Best Local Similarity:  99.86%         Mismatches:    1
Query Match:            84.39%         Indels:        0
DB:                     1              Gaps:          0

US-09-714-865-15 (1-2172) x DDX4_HUMAN (1-724)

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAspGluAspTrpGluAlaGluIleAsnProHisMetSerSerTyrValProIle 20

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 PheGluLysAspArgTyrSerGlyGluAsnGlyAspAsnPheAsnArgThrProAlaSer 40

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerSerGluMetAspAspGlyProSerArgArgAspHisPheMetLysSerGlyPheAla 60

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerGlyArgAsnPheGlyAsnArgAspAlaGlyGluCysAsnLysArgAspAsnThrSer 80

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrMetGlyGlyPheGlyValGlyLysSerPheGlyAsnArgGlyPheSerAsnSerArg 100

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheGluAspGlyAspSerSerGlyPheTrpArgGluSerSerAsnAspCysGluAspAsn 120

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ProThrArgAsnArgGlyPheSerLysArgGlyGlyTyrArgAspGlyAsnAsnSerGlu 140

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        141 AlaSerGlyProTyrArgArgGlyGlyArgGlySerPheArgGlyCysArgGlyGlyPhe 160

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GlyLeuGlySerProAsnAsnAspLeuAspProAspGluCysMetGlnArgThrGlyGly 180

Qy        541 CTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LeuPheGlySerArgArgProValLeuSerGlyThrGlyAsnGlyAspThrSerGlnSer 200

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ArgSerGlySerGlySerGluArgGlyGlyTyrLysGlyLeuAsnGluGluValIleThr 220

Qy        661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyGluSerSerAspThrGln 240

Qy        721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePheAla 260

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSerGly 280

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 HisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeuAsn 300

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIlePro 320

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 IleIleLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 340

Qy       1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaPheLeuLeuProIleLeuAlaHisMetMetHisAspGlyIleThrAlaSerArgPhe 360

Qy       1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuValAsnGln 380

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaValValIleTyr 400

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyGlyThrGlnLeuGlyHisSerIleArgGlnIleValGlnGlyCysAsnIleLeuCys 420

Qy       1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLysGlnIle 440

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlyProGluMet 460

Qy       1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 LysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeuMetPhe 480

Qy       1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SerAlaThrPheProGluGluIleGlnArgLeuAlaAlaGluPheLeuLysSerAsnTyr 500

Qy       1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThrValLeu 520

Qy       1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GlnValGlyGlnPheSerLysArgGluLysLeuValGluIleLeuArgAsnIleGlyAsp 540

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||   |||||||||||
Db        541 GluArgThrMetValPheValGluThrLysLysLysAlaAspPheIleAlaThrPheLeu 560

Qy       1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 CysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGluArgGlu 580

Qy       1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 GlnAlaLeuGlyAspPheArgPheGlyLysCysProValLeuValAlaThrSerValAla 600

Qy       1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAspLeuProSerThr 620

Qy       1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 IleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAla 640

Qy       1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 IleSerPhePheAspLeuGluSerAspAsnHisLeuAlaGlnProLeuValLysValLeu 660

Qy       1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ThrAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerThrTyrIle 680

Qy       2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 ProGlyPheSerGlySerThrArgGlyAsnValPheAlaSerValAspThrArgLysGly 700

Qy       2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCA 2136
              ||||||||||||||||||||||||||||||||||||
Db        701 LysSerThrLeuAsnThrAlaGlyPheSerSerSer 712
```

```
RESULT 2
DDX4_MOUSE
ID   DDX4_MOUSE     STANDARD;      PRT;   702 AA.
AC   Q61496; Q9D5X7;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 4 (VASA homolog) (Mvh).
GN   DDX4 OR VASA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
```

```
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE OF 60-702 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Testis;
RX   MEDLINE=95083681; PubMed=7991615;
RA   Fujiwara Y., Komiya T., Kawabata H., Sato M., Fujimoto H.,
RA   Furusawa M., Noce T.;
RT   "Isolation of a DEAD-family protein gene that encodes a murine
RT   homolog of Drosophila vasa and its specific expression in germ cell
RT   lineage.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:12258-12262(1994).
CC   -!- FUNCTION: May play a role in germ cell development.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND PERINUCLEAR.
CC   -!- TISSUE SPECIFICITY: Testis.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN SPERMATOGENIC CELLS FROM THE
CC       SPERMATOCYTE STAGE TO THE ROUND SPERMATID STAGE.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC       DDX4/VASA SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AK014844; BAB29578.1; -.
DR   EMBL; D14859; BAA03584.1; -.
DR   HSSP; Q58083; 1HV8.
DR   MGD; MGI:102670; Ddx4.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 2.
DR   Pfam; PF00271; helicase_C; 2.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.
FT   NP_BIND     305    312       ATP (POTENTIAL).
FT   SITE        419    422       DEAD BOX.
FT   DOMAIN       58    207       GLY-RICH.
FT   CONFLICT    152    152       R -> C (IN REF. 2).
FT   CONFLICT    156    160       LFGSR -> FLVLG (IN REF. 2).
FT   CONFLICT    277    277       A -> R (IN REF. 2).
FT   CONFLICT    291    291       S -> T (IN REF. 2).
FT   CONFLICT    368    369       RA -> IS (IN REF. 2).
FT   CONFLICT    423    423       R -> S (IN REF. 2).
FT   CONFLICT    430    430       G -> A (IN REF. 2).
FT   CONFLICT    433    433       M -> I (IN REF. 2).
FT   CONFLICT    448    448       R -> H (IN REF. 2).
FT   CONFLICT    472    472       S -> N (IN REF. 2).
FT   CONFLICT    499    509       SKREKLVEILR -> QKEKSLLRFYE (IN REF. 2).
FT   CONFLICT    540    540       T -> S (IN REF. 2).
FT   CONFLICT    680    702       HTLNTAGISSSQAPNPVDDESWD -> AHVEYSGDFFFTSS
FT                                QSS (IN REF. 2).
SQ   SEQUENCE   702 AA;  76470 MW;  5D1E4CC0BDF39748 CRC64;

Alignment Scores:
Pred. No.:              4.15e-74      Length:        702
Score:                  82.00         Matches:       82
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            11.33%        Indels:        0
DB:                     1             Gaps:          0

US-09-714-865-15 (1-2172) x DDX4_MOUSE (1-702)

Qy        718 CAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePhe 232

Qy        778 GCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSer 252

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeu 272

Qy        898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 AsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIle 292

Qy        958 CCTATC 963
              ||||||
Db        293 ProIle 294

RESULT 3
DDX4_RAT
ID   DDX4_RAT       STANDARD;      PRT;   713 AA.
AC   Q64060;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 4 (VASA homolog) (rVLG).
GN   DDX4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Testis;
RX   MEDLINE=95160706; PubMed=7857296;
RA   Komiya T., Tanigawa Y.;
RT   "Cloning of a gene of the DEAD box protein family which is
RT   specifically expressed in germ cells in rats.";
RL   Biochem. Biophys. Res. Commun. 207:405-410(1995).
CC   -!- FUNCTION: May play a role in germ cell development.
CC   -!- TISSUE SPECIFICITY: Testis.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC       DDX4/VASA SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; S75275; AAB33364.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.
FT   NP_BIND     317    324       ATP (POTENTIAL).
FT   SITE        431    434       DEAD BOX.
FT   DOMAIN       58    219       GLY-RICH.
SQ   SEQUENCE   713 AA;  77955 MW;  A31A4542EF7237F6 CRC64;

Alignment Scores:
Pred. No.:              4.14e-74      Length:        713
Score:                  82.00         Matches:       82
```

```
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             11.33%         Indels:        0
DB:                      1              Gaps:          0

US-09-714-865-15 (1-2172) x DDX4_RAT (1-713)

Qy        718 CAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePhe 244

Qy        778 GCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 AlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSer 264

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeu 284

Qy        898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 AsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIle 304

Qy        958 CCTATC 963
              ||||||
Db        305 ProIle 306

RESULT 4
VASA_DROME
ID   VASA_DROME     STANDARD;      PRT;   661 AA.
AC   P09052; Q24582; Q9V3Q8;
DT   01-NOV-1988 (Rel. 09, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Vasa protein (Antigen Mab46F11).
GN   VAS OR BG:DS00929.14 OR CG3506.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89014721; PubMed=3140040;
RA   Lasko P.F., Ashburner M.;
RT   "The product of the Drosophila gene vasa is very similar to
RT   eukaryotic initiation factor-4A.";
RL   Nature 335:611-617(1988).
RN   [2]
RP   REVISIONS.
RA   Lasko P.F.;
RL   Submitted (DEC-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89028669; PubMed=3052853;
RA   Hay B., Jan L.Y., Jan Y.N.;
RT   "A protein component of Drosophila polar granules is encoded by vasa
RT   and has extensive sequence similarity to ATP-dependent helicases.";
RL   Cell 55:577-587(1988).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=99403001; PubMed=10471707;
RA   Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA   Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA   Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA   Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA   Celniker S., Rubin G.M.;
RT   "An exploration of the sequence of a 2.9-Mb region of the genome of
RT   Drosophila melanogaster: the Adh region.";
RL   Genetics 153:179-219(1999).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: THE VASA PROTEIN IS REQUIRED ONLY IN THE FEMALE GERM
CC       LINE. IT IS IMPORTANT FOR OOCYTE FORMATION AND IN THE
CC       SPECIFICATION OF THE POSTERIOR STRUCTURES OF THE EMBRYO.
CC   -!- DEVELOPMENTAL STAGE: MATERNALLY EXPRESSED (DURING OOGENESIS).
CC       FUNCTION DURING EARLY EMBRYOGENESIS.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC       DDX4/VASA SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X12945; CAA31405.1; -.
DR   EMBL; X12946; CAA31405.1; JOINED.
DR   EMBL; M23560; AAA29013.1; -.
DR   EMBL; AE003412; AAF44917.1; -.
DR   EMBL; AE003646; AAF53438.1; -.
DR   PIR; A31922; A31922.
DR   PIR; S01676; S01676.
DR   HSSP; Q58083; 1HV8.
DR   FlyBase; FBgn0003970; vas.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
```

```
KW   Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.
FT   NP_BIND     289    296       ATP (BY SIMILARITY).
FT   SITE        399    402       DEAD BOX.
FT   DOMAIN       93    127       5 X 7 AA TANDEM REPEATS OF [FS]-R-G-G-
FT                                [EQ]-G-G.
FT   REPEAT       93     99       1.
FT   REPEAT      100    106       2.
FT   REPEAT      107    113       3.
FT   REPEAT      114    120       4.
FT   REPEAT      121    127       5.
FT   CONFLICT     35     35       A -> R (IN REF. 3).
FT   CONFLICT    153    165       MISSING (IN REF. 3).
FT   CONFLICT    192    192       V -> A (IN REF. 1 AND 3).
FT   CONFLICT    265    265       Y -> F (IN REF. 1).
FT   CONFLICT    322    322       V -> C (IN REF. 3).
FT   CONFLICT    452    452       F -> S (IN REF. 1).
FT   CONFLICT    582    582       R -> C (IN REF. 1).
FT   CONFLICT    594    594       D -> H (IN REF. 3).
SQ   SEQUENCE   661 AA;  72331 MW;  8617C25CCB3130B9 CRC64;

Alignment Scores:
Pred. No.:              1.82e-14      Length:        661
Score:                  23.00         Matches:       23
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.18%         Indels:        0
DB:                     1             Gaps:          0

US-09-714-865-15 (1-2172) x VASA_DROME (1-661)

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeu 301

Qy       1033 CCAATTTTG 1041
              |||||||||
Db        302 ProIleLeu 304

RESULT 5
PL10_MOUSE
ID   PL10_MOUSE     STANDARD;      PRT;   660 AA.
AC   P16381;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ATP-dependent RNA helicase PL10.
GN   D1PAS1 OR PL10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=89249320; PubMed=2720782;
RA   Leroy P., Alzari P., Sassoon D., Wolgemuth D., Fellous M.;
RT   "The protein encoded by a murine male germ cell-specific transcript
RT   is a putative ATP-dependent RNA helicase.";
RL   Cell 57:549-559(1989).
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. POSSIBLE ROLE
CC       IN A KEY STEP OF THE SPERMATOGENIC PROCESS.
CC   -!- TISSUE SPECIFICITY: TESTIS.
CC   -!- DEVELOPMENTAL STAGE: HIGH LEVELS OF PL10 DURING THE MEIOTIC AND
CC       HAPLOID STAGES OF SPERMATOGENESIS.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J04847; AAA39942.1; -.
DR   PIR; A32378; A32378.
DR   HSSP; Q58083; 1HV8.
DR   MGD; MGI:91842; D1Pas1.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; DNA-binding; Spermatogenesis.
FT   NP_BIND     223    230       ATP (POTENTIAL).
FT   SITE        346    349       DEAD BOX.
FT   DNA_BIND    495    514       POTENTIAL.
FT   DOMAIN      581    660       GLY/SER-RICH.
SQ   SEQUENCE   660 AA;  73140 MW;  50AD4E6A131AE603 CRC64;

Alignment Scores:
Pred. No.:              1.87e-13      Length:        660
Score:                  22.00         Matches:       22
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.04%         Indels:        0
DB:                     1             Gaps:          0

US-09-714-865-15 (1-2172) x PL10_MOUSE (1-660)

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 236

Qy       1036 ATTTTG 1041
              ||||||
Db        237 IleLeu 238

RESULT 6
DDX3_HUMAN
ID   DDX3_HUMAN     STANDARD;      PRT;   661 AA.
AC   O00571; O15536;
DT   15-JUL-1998 (Rel. 36, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 3 (Helicase-like protein 2) (HLP2) (DEAD-box, X
DE   isoform).
GN   DDX3 OR DBX.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver, and Hippocampus;
RA   Chung J., Lee S.-G., Song K.;
RT   "Identification of a human homolog of a putative RNA helicase gene
RT   (mDEAD3) expressed in mouse erythroid cells.";
RL   Korean J. Biochem. 27:193-197(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Owsianka A.M., Patel A.H.;
RT   "DEAD box putative RNA helicase.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98022381; PubMed=9381176;
RA   Lahn B.T., Page D.C.;
RT   "Functional coherence of the human Y chromosome.";
RL   Science 278:675-680(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Skin;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 1-9, AND ACETYLATION.
RX   MEDLINE=20318637; PubMed=10859333;
RA   Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,
RA   Lopez de Castro J.A.;
RT   "An N-acetylated natural ligand of human histocompatibility leukocyte
RT   antigen (HLA)-B39. Classical major histocompatibility complex class I
RT   proteins bind peptides with a blocked NH(2) terminus in vivo.";
RL   J. Exp. Med. 191:2083-2092(2000).
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. INTERACTS
CC       SPECIFICALLY WITH HEPATITIS C VIRUS CORE PROTEIN RESULTING A
CC       CHANGE IN INTRACELLULAR LOCATION.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U50553; AAB95637.1; -.
DR   EMBL; AF061337; AAC34298.1; -.
DR   EMBL; AF000983; AAC51830.1; -.
DR   EMBL; AF000982; AAC51829.1; -.
DR   EMBL; BC011819; AAH11819.1; -.
DR   HSSP; Q58083; 1HV8.
DR   Genew; HGNC:2745; DDX3.
DR   MIM; 300160; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; DNA-binding; Acetylation.
FT   INIT_MET      0      0
FT   NP_BIND     223    230       ATP (POTENTIAL).
FT   SITE        346    349       DEAD BOX.
FT   DOMAIN      581    661       GLY/SER-RICH.
FT   MOD_RES       1      1       ACETYLATION.
FT   CONFLICT     49     49       K -> R (IN REF. 3).
SQ   SEQUENCE   661 AA;  73112 MW;  F0F03DD8FBC00A65 CRC64;

Alignment Scores:
Pred. No.:               1.87e-13      Length:         661
Score:                   22.00         Matches:        22
Percent Similarity:      100.00%       Conservative:   0
Best Local Similarity:   100.00%       Mismatches:     0
Query Match:             3.04%         Indels:         0
DB:                      1             Gaps:           0

US-09-714-865-15 (1-2172) x DDX3_HUMAN (1-661)

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 236

Qy       1036 ATTTTG 1041
              ||||||
Db        237 IleLeu 238

RESULT 7
DDX3_MOUSE
ID   DDX3_MOUSE     STANDARD;      PRT;   661 AA.
AC   Q62167; O09060; O09143;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DEAD-box protein 3 (DEAD-box RNA helicase DEAD3) (mDEAD3) (Embryonic
DE   RNA helicase) (D1PAS1 related sequence 2).
GN   DDX3 OR DEAD3 OR ERH OR D1PAS1-RS2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6, and DBA;
RX   MEDLINE=97104282; PubMed=8948440;
RA   Sowden J.C., Putt W., Morrison K., Beddington R., Edwards Y.;
RT   "The embryonic RNA helicase gene (ERH): a new member of the DEAD box
RT   family of RNA helicases.";
RL   Biochem. J. 308:839-846(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythroleukemia;
RX   MEDLINE=94192995; PubMed=8144024;
RA   Gee S.L., Conboy J.G.;
RT   "Mouse erythroid cells express multiple putative RNA helicase genes
RT   exhibiting high sequence conservation from yeast to mammals.";
RL   Gene 140:171-177(1994).
RN   [3]
RP   SEQUENCE OF 1-9, AND ACETYLATION.
RX   MEDLINE=20318637; PubMed=10859333;
RA   Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,
RA   Lopez de Castro J.A.;
RT   "An N-acetylated natural ligand of human histocompatibility leukocyte
RT   antigen (HLA)-B39. Classical major histocompatibility complex class I
RT   proteins bind peptides with a blocked NH(2) terminus in vivo.";
RL   J. Exp. Med. 191:2083-2092(2000).
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. IT MAY PLAY A ROLE
CC       IN TRANSLATIONAL ACTIVATION OF MRNA IN THE OOCYTE AND EARLY
CC       EMBRYO.
CC   -!- TISSUE SPECIFICITY: DEVELOPMENTALLY REGULATED.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN OOCYTES. UBIQUITOUSLY FOUND IN 9
CC       DAYS POST-CONCEPTION EMBRYO, AT LATER STAGES IT IS RESTRICTED TO
CC       BRAIN AND KIDNEY.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z38117; CAA86261.1; -.
DR   EMBL; L25126; AAA53630.1; -.
DR   HSSP; Q58083; 1HV8.
DR   MGD; MGI:103064; Ddx3.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; DNA-binding; Acetylation.
FT   INIT_MET      0      0
FT   NP_BIND     223    230       ATP (POTENTIAL).
FT   SITE        346    349       DEAD BOX.
FT   DOMAIN      581    661       GLY/SER-RICH.
FT   DOMAIN      608    615       POLY-SER.
FT   DOMAIN      623    629       POLY-GLY.
FT   DOMAIN      632    640       POLY-GLY.
```

```
FT   MOD_RES       1      1       ACETYLATION.
SQ   SEQUENCE   661 AA;  72970 MW;  A1E1FAAB5D19F57B CRC64;

Alignment Scores:
Pred. No.:              1.87e-13       Length:        661
Score:                  22.00          Matches:       22
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.04%          Indels:        0
DB:                     1              Gaps:          0

US-09-714-865-15 (1-2172) x DDX3_MOUSE (1-661)

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 236

Qy       1036 ATTTTG 1041
              ||||||
Db        237 IleLeu 238

RESULT 8
AN3_XENLA
ID   AN3_XENLA      STANDARD;      PRT;   697 AA.
AC   P24346;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ATP-dependent RNA helicase An3.
GN   AN3.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91141586; PubMed=1996140;
RA   Gururajan R., Perry-O'Keefe H., Melton D.A., Weeks D.L.;
RT   "The Xenopus localized messenger RNA An3 may encode an ATP-dependent
RT   RNA helicase.";
RL   Nature 349:717-719(1991).
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X57328; CAA40605.1; -.
DR   PIR; S13654; S13654.
DR   PIR; S29676; S29676.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding.
FT   NP_BIND     265    272       ATP (BY SIMILARITY).
FT   SITE        388    391       DEAD BOX.
FT   DOMAIN      623    697       GLY/SER-RICH.
SQ   SEQUENCE   697 AA;  77302 MW;  F3DD23EB60B2E2EF CRC64;

Alignment Scores:
Pred. No.:              1.85e-13       Length:        697
Score:                  22.00          Matches:       22
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.04%          Indels:        0
DB:                     1              Gaps:          0

US-09-714-865-15 (1-2172) x AN3_XENLA (1-697)

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 278

Qy       1036 ATTTTG 1041
              ||||||
Db        279 IleLeu 280

RESULT 9
DDXY_HUMAN
ID   DDXY_HUMAN     STANDARD;      PRT;   660 AA.
AC   O15523;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   DEAD-box protein 3, Y-chromosomal.
GN   DBY.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98022381; PubMed=9381176;
RA   Lahn B.T., Page D.C.;
RT   "Functional coherence of the human Y chromosome.";
RL   Science 278:675-680(1997).
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF000985; AAC51832.1; -.
DR   EMBL; AF000984; AAC51831.1; -.
DR   HSSP; Q58083; 1HV8.
DR   Genew; HGNC:2699; DBY.
DR   MIM; 400010; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; DNA-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
FT   SITE        345    348       DEAD BOX.
SQ   SEQUENCE   660 AA;  73094 MW;  5A67DC082AEE9CFF CRC64;

Alignment Scores:
Pred. No.:              2.06e-09       Length:        660
Score:                  18.00          Matches:       18
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            2.49%          Indels:        0
DB:                     1              Gaps:          0
```

```
US-09-714-865-15 (1-2172) x DDXY_HUMAN (1-660)

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTG 1041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeu 237

RESULT 10
GLH1_CAEEL
ID   GLH1_CAEEL     STANDARD;      PRT;   763 AA.
AC   P34689; Q9TXH4;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ATP-dependent RNA helicase glh-1 (Germline helicase-1).
GN   GLH-1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94022363; PubMed=8415696;
RA   Roussell D.L., Bennett K.L.;
RT   "glh-1, a germ-line putative RNA helicase from Caenorhabditis, has
RT   four zinc fingers.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:9300-9304(1993).
RN   [2]
RP   REVISIONS TO 83-138; 275; 288 AND 398.
RA   Roussell D.L., McCrone J.S., Smith P.A., Gruidl M.E., Bennett K.L.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.
CC   -!- DEVELOPMENTAL STAGE: DURING GERM-LINE PROLIFERATION.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 4 CCHC-TYPE ZINC FINGERS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L19948; AAC27384.1; -.
DR   PIR; A48686; A48686.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00098; zf-CCHC; 4.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 4.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS50158; ZF_CCHC; 4.
KW   Helicase; ATP-binding; RNA-binding; Zinc-finger; Repeat.
FT   DOMAIN       24     93       7 X 10 AA TANDEM REPEATS, GLY-RICH.
FT   REPEAT       24     33       1.
FT   REPEAT       34     43       2.
FT   REPEAT       44     53       3.
FT   REPEAT       54     63       4.
FT   REPEAT       64     73       5.
FT   REPEAT       74     83       6.
FT   REPEAT       84     93       7.
FT   DOMAIN      207    236       GLY-RICH.
FT   ZN_FING     158    175       CCHC-TYPE 1.
FT   ZN_FING     183    200       CCHC-TYPE 2.
FT   ZN_FING     242    259       CCHC-TYPE 3.
FT   ZN_FING     262    279       CCHC-TYPE 4.
FT   NP_BIND     385    392       ATP (BY SIMILARITY).
FT   SITE        499    502       DEAD BOX.
SQ   SEQUENCE   763 AA;  79792 MW;  ADB69DE286A028D6 CRC64;

Alignment Scores:
Pred. No.:              2.02e-09      Length:        763
Score:                  18.00         Matches:       18
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.49%         Indels:        0
DB:                     1             Gaps:          0

US-09-714-865-15 (1-2172) x GLH1_CAEEL (1-763)

Qy        985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 MetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIle 399

RESULT 11
DBP1_YEAST
ID   DBP1_YEAST     STANDARD;      PRT;   617 AA.
AC   P24784; P20446;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable ATP-dependent RNA helicase DBP1 (Helicase CA1).
GN   DBP1 OR YPL119C OR LPH8C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBY939;
RX   MEDLINE=91312117; PubMed=1857205;
RA   Jamieson D.J., Beggs J.D.;
RT   "A suppressor of yeast spp81/ded1 mutations encodes a very similar
RT   putative ATP-dependent RNA helicase.";
RL   Mol. Microbiol. 5:805-812(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313271; PubMed=9169875;
RA   Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
RA   Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA   Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA   Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA   Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA   Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA   Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA   Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA   Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA   Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA   Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA   Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA   Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA   Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA   Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA   Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL   Nature 387:103-105(1997).
RN   [3]
RP   SEQUENCE OF 316-500 FROM N.A.
RX   MEDLINE=90160368; PubMed=2406722;
RA   Chang T.-H., Arenas J., Abelson J.;
RT   "Identification of five putative yeast RNA helicase genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:1571-1575(1990).
CC   -!- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE. SUPPRESSOR OF A
CC       SPP81/DED1 MUTATION.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X55993; CAA39465.1; -.
DR   EMBL; U43503; AAB68243.1; -.
DR   PIR; S16790; S16790.
DR   PIR; A34848; A34848.
DR   HSSP; Q58083; 1HV8.
DR   SGD; S0006040; DBP1.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding.
FT   NP_BIND     198    205       ATP (BY SIMILARITY).
FT   SITE        318    321       DEAD BOX.
FT   CONFLICT     43     44       ST -> RS (IN REF. 1).
FT   CONFLICT     48     48       E -> K (IN REF. 1).
FT   CONFLICT     88     88       G -> R (IN REF. 1).
FT   CONFLICT    115    115       E -> QK (IN REF. 1).
FT   CONFLICT    496    496       V -> I (IN REF. 3).
SQ   SEQUENCE   617 AA;  67917 MW;  8152404B7628671D CRC64;

Alignment Scores:
Pred. No.:              2.14e-08       Length:        617
Score:                  17.00          Matches:       17
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            2.35%          Indels:        0
DB:                     1              Gaps:          0

US-09-714-865-15 (1-2172) x DBP1_YEAST (1-617)

Qy       1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 IleLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 328

RESULT 12
GLH3_CAEEL
ID   GLH3_CAEEL     STANDARD;      PRT;   720 AA.
AC   O01836;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ATP-dependent RNA helicase glh-3 (Germline helicase-3).
GN   GLH-3 OR B0414.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=20311358; PubMed=10851135;
RA   Kuznicki K.A., Smith P.A., Leung-Chiu W.M., Estevez A.O., Scott H.C.,
RA   Bennett K.L.;
RT   "Combinatorial RNA interference indicates GLH-4 can compensate for
RT   GLH-1; these two P granule components are critical for fertility in
RT   C. elegans.";
RL   Development 127:2907-2916(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Sammons L., Wohldmann P., Rohlfing T.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
```

```
CC   -!- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.
CC   -!- DEVELOPMENTAL STAGE: DURING GERM-LINE PROLIFERATION.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF079509; AAC28388.1; -.
DR   EMBL; AF003145; AAB57719.1; -.
DR   HSSP; Q58083; 1HV8.
DR   WormPep; B0414.6; CE07736.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00098; zf-CCHC; 2.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 2.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS50158; ZF_CCHC; 2.
KW   Helicase; ATP-binding; RNA-binding; Zinc-finger; Repeat.
FT   ZN_FING     202    219       CCHC-TYPE 1.
FT   ZN_FING     222    239       CCHC-TYPE 2.
FT   NP_BIND     342    349       ATP (BY SIMILARITY).
FT   SITE        456    459       DEAD BOX.
SQ   SEQUENCE   720 AA;  79728 MW;  A91F49FE2B79DEAA CRC64;

Alignment Scores:
Pred. No.:              2.09e-08       Length:        720
Score:                  17.00          Matches:       17
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            2.35%          Indels:        0
DB:                     1              Gaps:          0

US-09-714-865-15 (1-2172) x GLH3_CAEEL (1-720)

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIle 356

RESULT 13
DEA2_MOUSE
ID   DEA2_MOUSE     STANDARD;      PRT;   245 AA.
AC   Q62095;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DEAD-box RNA helicase DEAD2 (mDEAD2) (D1PAS1 related sequence 1)
DE   (Fragment).
GN   DEAD2 OR D1PAS1-RS1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythroleukemia;
RX   MEDLINE=94192995; PubMed=8144024;
RA   Gee S.L., Conboy J.G.;
RT   "Mouse erythroid cells express multiple putative RNA helicase genes
RT   exhibiting high sequence conservation from yeast to mammals.";
RL   Gene 140:171-177(1994).
```

```
CC   -!- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. IT MAY PLAY A ROLE
CC       IN TRANSLATIONAL ACTIVATION OF MRNA IN THE OOCYTE AND EARLY
CC       EMBRYO.
CC   -!- TISSUE SPECIFICITY: FOUND IN HEART, BRAIN, LIVER, SKELETAL MUSCLE,
CC       AND KIDNEY. LOW EXPRESSION WAS DETECTED IN LUNG.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L25337; AAA53631.1; -.
DR   HSSP; Q58083; 1HV8.
DR   MGD; MGI:99524; D1Pas1-rs1.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding.
FT   NON_TER       1      1
FT   SITE         64     67       DEAD BOX.
FT   NON_TER     245    245
SQ   SEQUENCE   245 AA;  27978 MW;  276FC43820B3EDDA CRC64;

Alignment Scores:
Pred. No.:              2.52e-07        Length:         245
Score:                  16.00           Matches:        16
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.21%           Indels:         0
DB:                     1               Gaps:           0

US-09-714-865-15 (1-2172) x DEA2_MOUSE (1-245)

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 74

RESULT 14
DED1_YEAST
ID   DED1_YEAST     STANDARD;      PRT;   604 AA.
AC   P06634;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable ATP-dependent RNA helicase DED1.
GN   DED1 OR SPP81 OR YOR204W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A364A X H79-20.3;
RX   MEDLINE=91141585; PubMed=1996139;
RA   Jamieson D.J., Rahe B., Pringle J., Beggs J.D.;
RT   "A suppressor of a yeast splicing mutation (prp8-1) encodes a
RT   putative ATP-dependent RNA helicase.";
RL   Nature 349:715-717(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Hughes B., Pohl T.M.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-112 FROM N.A.
RX   MEDLINE=86093663; PubMed=3001645;
RA   Struhl K.;
RT   "Nucleotide sequence and transcriptional mapping of the yeast
RT   pet56-his3-ded1 gene region.";
RL   Nucleic Acids Res. 13:8587-8601(1985).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=97289744; PubMed=9144215;
RA   de la Cruz J., Iost I., Kressler D., Linder P.;
RT   "The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent
RT   translation in Saccharomyces cerevisiae.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:5201-5206(1997).
CC   -!- FUNCTION: POSITIVELY INVOLVED IN THE INITIATION OF PROTEIN
CC       SYNTHESIS. PROBABLE ATP-BINDING RNA HELICASE. ESSENTIAL FOR CELL
CC       VIABILITY. SUPPRESSOR OF A PRP8 MUTATION. MAY PLAY A ROLE IN MRNA
CC       SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X57278; CAA40546.1; -.
DR   EMBL; Z75110; CAA99419.1; -.
DR   EMBL; X03245; CAA27004.1; -.
DR   PIR; S13653; S13653.
DR   PIR; S07683; S07683.
DR   HSSP; Q58083; 1HV8.
DR   SGD; S0005730; DED1.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding.
FT   NP_BIND     186    193       ATP (BY SIMILARITY).
FT   SITE        306    309       DEAD BOX.
FT   CONFLICT     37     37       S -> M (IN REF. 2).
SQ   SEQUENCE   604 AA;  65553 MW;  B6722D94C03BFA4B CRC64;

Alignment Scores:
Pred. No.:              2.2e-07         Length:         604
Score:                  16.00           Matches:        16
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.21%           Indels:         0
DB:                     1               Gaps:           0

US-09-714-865-15 (1-2172) x DED1_YEAST (1-604)

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 316

RESULT 15
DED1_SCHPO
ID   DED1_SCHPO     STANDARD;      PRT;   636 AA.
AC   O13370; O59857;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   ATP-dependent RNA helicase ded1.
GN   DED1 OR SUM3 OR DEP1 OR MOC2 OR SPCC1795.11.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
```

```
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, AND ALTERNATIVE FORMS.
RX   MEDLINE=21659741; PubMed=11711540;
RA   Liu H.-Y., Nefsky B.S., Walworth N.C.;
RT   "The ded1 DEAD box helicase interacts with Chk1 and Cdc2.";
RL   J. Biol. Chem. 277:2637-2643(2002).
RN   [2]
RP   SEQUENCE FROM N.A., AND FUNCTION.
RX   MEDLINE=20191909; PubMed=10725227;
RA   Grallert B., Kearsey S.E., Lenhard M., Carlson C.R., Nurse P.,
RA   Boye E., Labib K.;
RT   "A fission yeast general translation factor reveals links between
RT   protein synthesis and cell cycle controls.";
RL   J. Cell Sci. 113:1447-1458(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99326140; PubMed=10395922;
RA   Kawamukai M.;
RT   "Isolation of a novel gene, moc2, encoding a putative RNA helicase as
RT   a suppressor of sterile strains in Schizosaccharomyces pombe.";
RL   Biochim. Biophys. Acta 1446:93-101(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99054689; PubMed=9832516;
RA   Forbes K.C., Humphrey T., Enoch T.;
RT   "Suppressors of cdc25p overexpression identify two pathways that
RT   influence the G2/M checkpoint in fission yeast.";
RL   Genetics 150:1361-1375(1998).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- FUNCTION: Positively involved in the initiation of protein
CC       synthesis. Probable ATP-binding RNA helicase. Essential for cell
CC       viability. May play a role in mRNA splicing. Inactivation of ded1
CC       blocks mitotic cell cycle progression at G1 and G2/M.
CC   -!- SUBUNIT: Interacts with chk1, which is required for cell cycle
CC       arrest following DNA damage.
CC   -!- MISCELLANEOUS: A different form of ded1 has been identified via
CC       SDS-PAGE studies. It is uncertain how this form arises in vivo.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF025536; AAC04893.1; -.
DR   EMBL; AF084222; AAC34121.1; -.
DR   EMBL; AB012389; BAA25324.1; -.
DR   EMBL; AJ237697; CAB40192.1; -.
DR   EMBL; AL022598; CAA18646.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_ATP_helicase.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding; RNA-binding; Cell cycle; Mitosis.
FT   NP_BIND     213    220       ATP (POTENTIAL).
FT   SITE        336    339       DEAD BOX.
FT   CONFLICT     41     41       A -> R (IN REF. 3).
SQ   SEQUENCE   636 AA;  69758 MW;  094630A41A3C26F1 CRC64;

Alignment Scores:
Pred. No.:              2.18e-07      Length:         636
Score:                  16.00         Matches:        16
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            2.21%         Indels:         0
DB:                     1             Gaps:           0

US-09-714-865-15 (1-2172) x DED1_SCHPO (1-636)

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 221

Search completed: June 10, 2003, 17:05:53
Job time : 53.5 secs
```

GenCore version 5.1.6


```
OM nucleic - protein search, using frame_plus_n2p model

Run on:         June 10, 2003, 16:57:41 ; Search time 38 Seconds
                                           (without alignments)
                                           10989.682 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  724
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  OLIGO
                Xgapop 60.0 , Xgapext 60.0
                Ygapop 60.0 , Ygapext 60.0
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       283224 seqs, 96134422 residues

Word size:      1

Total number of hits satisfying chosen parameters:      565536

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111833_26595/app_query.fasta_1.2311
-DB=PIR_73 -QFMT=fastan -SUFFIX=oligo.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_55_@runat_05062003_111833_26595 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :      PIR_73:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	522	72.1	635	2	T46407	probable RNA helic
2	82	11.3	713	2	JC2534	RVLG protein - rat
3	64	8.8	637	2	I49638	probable RNA helic
4	24	3.3	700	2	I51235	DEAD box protein -
5	23	3.2	661	2	A58768	ATP-dependent RNA
6	22	3.0	660	2	A32378	probable ATP-depen
7	22	3.0	662	1	I84741	RNA helicase - mou
8	22	3.0	697	1	S13654	ATP-dependent RNA
9	19	2.6	646	2	T45677	ATP-dependent RNA
10	18	2.5	604	2	C87818	protein glh-1 [imp
11	18	2.5	604	2	T15132	ATP-dependent RNA
12	18	2.5	633	2	H84854	probable ATP-depen
13	17	2.3	603	2	T45671	ATP-dependent RNA
14	17	2.3	617	2	S62003	probable ATP-depen
15	17	2.3	688	2	T48796	probable ATP-depen
16	17	2.3	720	2	T15231	germline RNA helic
17	16	2.2	604	2	S13653	ATP-dependent RNA
18	16	2.2	636	2	T43543	probable ATP-depen
19	16	2.2	707	2	A48686	probable RNA helic
20	15	2.1	263	2	T51349	RNA helicase RH30
21	15	2.1	501	2	B96593	probable ethylene-
22	15	2.1	546	1	S13757	RNA helicase DBP2
23	15	2.1	550	1	S14048	RNA helicase dbp2
24	15	2.1	561	2	T22917	probable ATP-depen
25	15	2.1	575	1	S11485	RNA helicase - fru
26	15	2.1	607	1	S42639	ATP-dependent RNA
27	15	2.1	614	1	I48385	RNA helicase TNZ2
28	15	2.1	614	2	JC1087	RNA helicase, ATP-
29	15	2.1	619	2	T52137	ATP-dependent DEAD
30	15	2.1	650	2	S72367	ATP-dependent RNA
31	15	2.1	713	2	T48634	DRH1 DEAD box prot
32	14	1.9	422	2	H82418	ATP-dependent RNA
33	14	1.9	423	2	F82295	ATP-dependent RNA
34	14	1.9	425	2	T46969	DEAD box RNA helic
35	14	1.9	426	2	AF2395	ATP-dependent RNA
36	14	1.9	446	2	E82835	ATP-dependent RNA
37	14	1.9	446	2	A83213	probable ATP-depen
38	14	1.9	451	2	E82488	ATP-dependent RNA
39	14	1.9	454	2	E64816	probable ATP-depen
40	14	1.9	455	2	C90738	probable ATP-depen
41	14	1.9	455	2	D85588	probable ATP-depen
42	14	1.9	455	2	AI0599	probable ATP-depen
43	14	1.9	458	2	E83588	RNA helicase DbpA
44	14	1.9	460	2	AF0216	ATP-dependent RNA
45	14	1.9	462	2	E81090	ATP-dependent RNA
46	14	1.9	462	2	G81850	probable ATP-depen
47	14	1.9	535	2	AH3368	ATP-dependent RNA
48	14	1.9	639	2	D83591	probable ATP-depen
49	13	1.8	368	2	F69531	ATP-dependent RNA
50	13	1.8	397	2	C83163	ATP-dependent RNA
51	13	1.8	397	2	D82203	ATP-dependent RNA
52	13	1.8	411	2	T51739	RNA helicase RH5 [
53	13	1.8	457	2	F81857	probable ATP-depen
54	13	1.8	457	2	H81085	ATP-dependent RNA
55	13	1.8	504	2	T27176	probable ATP-depen
56	13	1.8	537	2	A86444	probable RNA helic
57	12	1.7	441	2	AD0330	ATP-dependent RNA
58	12	1.7	444	2	AD0830	ATP-dependent RNA
59	12	1.7	444	2	B91059	ATP-dependent RNA
60	12	1.7	444	2	G65035	ATP-dependent RNA
61	12	1.7	542	2	T05988	hypothetical prote
62	12	1.7	578	2	S67386	probable ATP-depen
63	12	1.7	701	2	T15942	hypothetical prote
64	12	1.7	1032	2	A57514	RNA helicase HEL11
65	11	1.5	343	2	A99353	ATP-dependent RNA
66	11	1.5	360	2	G90370	ATP-dependent RNA
67	11	1.5	444	2	H85903	ATP-dependent RNA
68	11	1.5	449	2	F83151	probable ATP-depen
69	11	1.5	452	2	C82506	ATP-dependent RNA
70	11	1.5	457	2	G90869	ATP-dependent RNA
71	11	1.5	457	2	B85749	ATP-dependent RNA
72	11	1.5	457	2	B64884	ATP-dependent RNA
73	11	1.5	482	2	AE3381	ATP-dependent RNA
74	11	1.5	491	2	G86297	F3O9.8 protein - A
75	11	1.5	503	2	A97581	ATP-dependent RNA
76	11	1.5	503	2	AG2801	ATP-dependent RNA
77	11	1.5	517	2	B87478	hypothetical prote
78	11	1.5	523	1	S30805	probable RNA helic
79	11	1.5	543	2	S46713	ATP-dependent RNA
80	11	1.5	601	2	D84973	ATP-dependent RNA
81	11	1.5	613	2	F64056	probable ATP-depen
82	11	1.5	646	2	C91134	inducible ATP-inde
83	11	1.5	646	2	F85979	inducible ATP-inde
84	11	1.5	646	2	AB0901	ATP-dependent RNA
85	11	1.5	646	2	F65106	probable ATP-depen
86	11	1.5	659	2	JX0314	DEAD box protein -
87	11	1.5	662	2	T41512	u5 snrnp-like RNA

	88	11	1.5	664	2	AI0423	cold-shock dead-bo
	89	11	1.5	1032	2	G89427	protein T08D2.3 [i
	90	11	1.5	1156	2	T43326	germline RNA helic
	91	11	1.5	1172	2	T32759	hypothetical prote
	92	10	1.4	187	2	T51345	RNA helicase RH20
	93	10	1.4	224	2	T51742	RNA helicase RH11
	94	10	1.4	240	2	T46269	hypothetical prote
	95	10	1.4	376	2	T51342	RNA helicase RH10
	96	10	1.4	418	2	E64100	probable ATP-depen
	97	10	1.4	421	2	A91218	probable ATP-depen
	98	10	1.4	421	2	B86064	probable ATP-depen
	99	10	1.4	421	2	AG0922	probable ATP-depen
	100	10	1.4	421	2	G65181	rhlB protein - Esc
	101	10	1.4	425	1	H69124	translation initia
	102	10	1.4	428	2	AE0471	probable DEAD-box
	103	10	1.4	438	2	C82340	ATP-dependent RNA
	104	10	1.4	460	2	B82060	ATP-dependent RNA
	105	10	1.4	470	2	AE1584	ATP-dependent RNA
	106	10	1.4	470	2	AF1230	ATP-dependent RNA
	107	10	1.4	502	2	C97640	probable ATP-depen
	108	10	1.4	502	2	AD2863	dead-box ATP-depen
	109	10	1.4	503	2	F95988	probable ATP-depen
	110	10	1.4	511	2	D69772	ATP-dependent RNA
	111	10	1.4	513	2	AD1959	ATP-dependent RNA
	112	10	1.4	516	2	AC1540	ATP-dependent RNA
	113	10	1.4	520	2	AB1183	ATP-dependent RNA
	114	10	1.4	524	2	G95184	ATP-dependent RNA
	115	10	1.4	524	2	G98051	conserved hypothet
	116	10	1.4	527	2	D70595	probable ATP-depen
	117	10	1.4	528	2	D97270	ATP-dependent RNA
	118	10	1.4	539	2	H83947	ATP-dependent RNA
	119	10	1.4	544	2	E87010	probable ATP-depen
	120	10	1.4	547	2	F86668	ATP-dependent RNA
	121	10	1.4	563	2	E70752	probable deaD prot
	122	10	1.4	564	2	S59649	ATP-dependent RNA
	123	10	1.4	567	2	A83292	probable ATP-depen
	124	10	1.4	610	2	G86407	hypothetical prote
	125	10	1.4	614	2	G82830	ATP-dependent RNA
	126	10	1.4	646	2	H88637	protein F53H1.1 [i
	127	10	1.4	748	2	H84913	probable ATP-depen
	128	10	1.4	845	2	D96799	hypothetical prote
	129	10	1.4	1166	2	H86341	hypothetical prote
c	130	9	1.3	162	2	C85356	glycine-rich prote
c	131	9	1.3	168	2	A03864	hypothetical 17.7K
	132	9	1.2	215	2	T49743	probable rrm-type
c	133	9	1.3	228	2	T01350	hypothetical prote
c	134	9	1.3	323	2	JC4774	fiber protein - hu
	135	9	1.2	376	2	C69813	RNA helicase homol
c	136	9	1.3	393	2	T49578	hypothetical prote
	137	9	1.2	431	2	S48908	helicase homolog -
	138	9	1.2	439	2	H64066	ATP-dependent RNA
	139	9	1.2	439	2	JC2118	DEAD-box RNA helic
	140	9	1.2	441	2	AB1290	ATP-dependent RNA
	141	9	1.2	442	2	AH1661	ATP-dependent RNA
c	142	9	1.3	453	2	B48210	ezy-1 protein prec
	143	9	1.2	481	2	S62423	ATP-dependent RNA
	144	9	1.2	481	2	C97238	ATP dependent RNA
c	145	9	1.3	508	2	A32225	nerve growth facto
c	146	9	1.3	543	2	A41211	early growth respo
	147	9	1.2	582	2	S53814	DEAD box protein -
	148	9	1.2	606	2	S14942	RNA helicase SPB4
	149	9	1.2	733	2	H84748	hypothetical prote
	150	9	1.2	746	2	T29584	hypothetical prote
	151	9	1.2	752	2	S64750	probable ATP-depen
	152	9	1.2	754	2	S62561	ATP dependent RNA
c	153	9	1.3	788	2	JS0747	regulatory protein
	154	9	1.2	789	2	T51310	RNA helicase RH28
c	155	9	1.3	809	2	S55344	outer envelope mem
	156	9	1.2	811	2	T19974	hypothetical prote
	157	9	1.2	819	2	T08745	probable RNA helic
	158	9	1.2	848	2	S59645	probable atp-depen
	159	9	1.2	995	2	S67564	hypothetical prote
	160	9	1.2	1212	2	C82834	DNA polymerase III

	161	8	1.1	71	2	T28271	ORF MSV109 hypothe
c	162	8	1.1	83	2	S24211	Ig kappa chain V r
c	163	8	1.1	91	2	S42186	Ig kappa chain V r
	164	8	1.1	92	2	A31409	actin - California
c	165	8	1.1	100	2	F90387	hypothetical prote
c	166	8	1.1	101	2	A33730	Ig kappa chain V r
c	167	8	1.1	112	2	A36259	ig kappa chain V r
c	168	8	1.1	113	2	F30560	Ig kappa chain V r
c	169	8	1.1	115	1	K2HUCM	Ig kappa chain V-I
c	170	8	1.1	121	2	S24205	Ig kappa chain V r
	171	8	1.1	122	2	AD2072	hypothetical prote
c	172	8	1.1	129	2	G84768	hypothetical prote
c	173	8	1.1	130	2	S40321	Ig kappa chain - h
c	174	8	1.1	132	2	C32513	Ig kappa chain pre
	175	8	1.1	143	2	E64602	ribonuclease H - H
c	176	8	1.1	159	2	T31598	hypothetical prote
c	177	8	1.1	174	2	T29783	hypothetical prote
c	178	8	1.1	188	2	T02247	hypothetical prote
	179	8	1.1	259	2	T51745	RNA helicase RH16
c	180	8	1.1	265	2	S46534	ubiquinol-cytochro
c	181	8	1.1	267	1	CTPGP	corticotropin / li
c	182	8	1.1	267	2	AD1835	hypothetical prote
c	183	8	1.1	268	2	T51678	myb-related transc
c	184	8	1.1	271	2	G01790	probable thiol-spe
	185	8	1.1	274	2	B70695	probable enoyl-coA
c	186	8	1.1	290	2	T05009	hypothetical prote
c	187	8	1.1	296	2	A47318	RNA-binding protei
c	188	8	1.1	296	2	I53142	gene Merc protein
c	189	8	1.1	297	2	C48210	ezy-1 protein (clo
	190	8	1.1	304	2	D84652	probable casein ki
c	191	8	1.1	320	2	C85440	myb-related protei
c	192	8	1.1	320	2	S09208	chorion protein s3
	193	8	1.1	323	2	C89901	riboflavin kinase
	194	8	1.1	323	2	AH2139	hypothetical prote
	195	8	1.1	328	2	B49543	maltosaccharide ut
	196	8	1.1	328	2	B95247	maltose operon tra
	197	8	1.1	328	2	G98111	maltose operon tra
c	198	8	1.1	341	2	B75374	conserved hypothet
	199	8	1.1	345	2	C95873	probable malate de
	200	8	1.1	345	2	AH3131	malate dehydrogena
	201	8	1.1	345	2	AD3635	malate dehydrogena
	202	8	1.1	346	2	C98156	probable L-malate
c	203	8	1.1	351	2	C82755	conserved hypothet
	204	8	1.1	360	2	T51344	RNA helicase RH18
	205	8	1.1	367	1	E64383	translation initia
	206	8	1.1	374	2	E96996	ATP-dependent RNA
c	207	8	1.1	379	2	E88109	protein T24E12.8 (
	208	8	1.1	381	2	H90136	RNA helicase [impo
c	209	8	1.1	382	2	S09139	coII intron protei
	210	8	1.1	399	2	T46259	hypothetical prote
	211	8	1.1	407	2	A71673	probable ATP-depen
	212	8	1.1	414	2	D97827	ATP-dependent RNA
c	213	8	1.1	415	2	D96759	probable serine ca
c	214	8	1.1	423	2	G96554	hypothetical prote
	215	8	1.1	428	2	H82390	ATP-dependent RNA
c	216	8	1.1	438	2	B96759	protein serine car
c	217	8	1.1	441	2	C96759	protein serine car
	218	8	1.1	451	2	T51744	RNA helicase RH13
c	219	8	1.1	453	2	S74698	UDP-glucose dehydr
c	220	8	1.1	462	4	S33798	FUS/CHOP mutant fu
c	221	8	1.1	464	2	A56600	intermediate filam
	222	8	1.1	465	2	E96737	probable DEAD/DEAH
c	223	8	1.1	469	1	I46076	DNA-directed DNA p
	224	8	1.1	474	2	E85175	ATP-dependent RNA
	225	8	1.1	476	2	H87352	hypothetical prote
	226	8	1.1	477	2	T01202	probable RNA helic
	227	8	1.1	478	2	I49731	RNA helicase - mou
	228	8	1.1	479	2	E69613	ATP-dependent RNA
	229	8	1.1	480	2	E90520	atp-dependent rna
	230	8	1.1	482	2	S66920	probable RNA helic
	231	8	1.1	489	2	S40731	ATP-dependent RNA
	232	8	1.1	491	2	B71957	ATP-dependent RNA
	233	8	1.1	492	2	G64550	ATP-dependent RNA

	234	8	1.1	492	2	S76612	ATP-dependent RNA
c	235	8	1.1	496	2	F75257	hypothetical prote
	236	8	1.1	496	2	T04020	hypothetical prote
	237	8	1.1	498	2	T51743	RNA helicase RH12
	238	8	1.1	498	2	T47916	DEAD box RNA helic
	239	8	1.1	499	2	T01230	probable RNA helic
	240	8	1.1	503	2	T39375	probable ATP-depen
	241	8	1.1	505	2	T51741	RNA helicase RH8 [
	242	8	1.1	505	2	S47451	hypothetical prote
	243	8	1.1	506	2	G90000	hypothetical prote
c	244	8	1.1	526	1	S33799	RNA-binding protei
	245	8	1.1	527	2	S38329	probable ATP-depen
	246	8	1.1	528	2	T02466	probable ATP-depen
	247	8	1.1	544	2	D82526	ATP-dependent RNA
	248	8	1.1	546	2	T14604	hypothetical 59.6K
	249	8	1.1	553	2	D96684	hypothetical prote
	250	8	1.1	566	2	S53813	RNA helicase - sli
c	251	8	1.1	569	1	KRMSE1	keratin, 59K type
c	252	8	1.1	570	2	S07330	keratin, epidermal
	253	8	1.1	571	2	T33113	hypothetical prote
	254	8	1.1	572	2	G02845	zyxin - human
	255	8	1.1	578	2	S62574	probable ATP-depen
	256	8	1.1	588	2	A39624	probable helicase
	257	8	1.1	595	2	T41007	atp dependent heli
	258	8	1.1	604	2	T41249	DEAD box ATP-depen
	259	8	1.1	606	2	T39966	probable atp depen
	260	8	1.1	610	2	S71758	DEAD box protein M
	261	8	1.1	613	2	T24662	hypothetical prote
	262	8	1.1	615	2	F98217	inducible ATP-inde
	263	8	1.1	615	2	AE3069	cold-shock dead-bo
	264	8	1.1	626	2	T10237	RNA helicase RH16
	265	8	1.1	626	2	AB3480	ATP-dependent RNA
	266	8	1.1	633	2	S69734	hypothetical prote
	267	8	1.1	638	2	T24661	hypothetical prote
c	268	8	1.1	643	1	KRHU2	keratin 1, type II
	269	8	1.1	648	2	T40504	atp-dependent, dea
c	270	8	1.1	659	2	T02838	probable membrane
	271	8	1.1	663	2	C82415	ATP-dependent RNA
	272	8	1.1	683	2	D71433	hypothetical prote
	273	8	1.1	685	2	T09159	RNA helicase prh75
	274	8	1.1	709	2	T39930	probable atp-depen
	275	8	1.1	742	2	S38093	probable purine nu
	276	8	1.1	748	2	T51738	RNA helicase RH3 [
	277	8	1.1	752	2	D95246	glycogen phosphory
	278	8	1.1	752	2	A98111	maltodextrin phosp
	279	8	1.1	770	2	S56805	probable RNA helic
	280	8	1.1	773	2	S46011	probable pre-mRNA-
c	281	8	1.1	787	2	C84898	hypothetical prote
c	282	8	1.1	801	2	A37353	membrane protein 4
	283	8	1.1	801	2	PC6010	RNA helicase Gu -
c	284	8	1.1	805	2	A56199	transcription fact
c	285	8	1.1	810	2	T48835	lethal(2)denticlel
c	286	8	1.1	823	2	A36378	probable transcrip
	287	8	1.1	853	2	A95269	probable sensory t
c	288	8	1.1	854	1	WMBP12	gene 12 protein -
c	289	8	1.1	858	2	A46613	protein 4.1, P4.1
c	290	8	1.1	863	2	H69527	valyl-tRNA synthet
c	291	8	1.1	873	2	B53225	ecdysone-induced p
	292	8	1.1	920	2	T41050	conserved hypothet
	293	8	1.1	971	2	S50912	probable membrane
	294	8	1.1	1022	2	T24663	hypothetical prote
c	295	8	1.1	1295	2	A32901	glp1 protein precu
c	296	8	1.1	1585	2	T31611	hypothetical prote
	297	8	1.1	1660	2	A70869	hypothetical glyci
c	298	8	1.1	2083	2	T42721	CRP-ductin-alpha p
c	299	8	1.1	2142	2	B35098	MHC class III hist
	300	8	1.1	2175	1	S03170	homeotic protein c
c	301	7	1.0	28	2	S58389	T-cell receptor be
c	302	7	1.0	42	2	D85604	hypothetical prote
	303	7	1.0	45	2	H70249	hypothetical prote
c	304	7	1.0	51	2	S64676	acetylcholinestera
	305	7	1.0	63	2	G85624	hypothetical prote
	306	7	1.0	63	2	B90761	hypothetical prote

c	307	7	1.0	63	2	T12121	NADH dehydrogenase
c	308	7	1.0	64	2	A64247	ribosomal protein
c	309	7	1.0	64	2	E87083	50S ribosomal prot
	310	7	1.0	65	2	T25646	hypothetical prote
c	311	7	1.0	65	2	B89948	hypothetical prote
c	312	7	1.0	68	2	T29417	hypothetical prote
c	313	7	1.0	72	2	I53411	acetylcholinestera
c	314	7	1.0	73	2	S78280	conserved hypothet
c	315	7	1.0	73	2	AH2788	hypothetical prote
c	316	7	1.0	83	2	A97568	hypothetical prote
c	317	7	1.0	89	2	B25155	Ig kappa chain V r
	318	7	1.0	90	2	G71368	conserved hypothet
c	319	7	1.0	91	1	W5WL13	E5 protein - human
c	320	7	1.0	91	2	F81996	hypothetical prote
	321	7	1.0	92	2	F71868	hypothetical prote
	322	7	1.0	95	2	F41298	kinesin-like prote
	323	7	1.0	96	2	A57483	3-mercaptopyruvate
	324	7	1.0	98	2	S53815	DEAD box protein -
c	325	7	1.0	99	2	AB2242	hypothetical prote
	326	7	1.0	104	1	G69482	hypothetical prote
c	327	7	1.0	104	2	E69767	transcription regu
	328	7	1.0	105	2	B90021	50S ribosomal prot
	329	7	1.0	109	2	G87661	hypothetical prote
c	330	7	1.0	110	2	F84670	hypothetical prote
c	331	7	1.0	112	2	D90193	conserved hypothet
c	332	7	1.0	112	4	S59333	hypothetical prote
	333	7	1.0	116	2	S57106	probable membrane
c	334	7	1.0	116	2	A72685	hypothetical prote
c	335	7	1.0	117	2	T11933	NADH2 dehydrogenas
	336	7	1.0	118	2	T16510	hypothetical prote
c	337	7	1.0	120	2	T19392	hypothetical prote
c	338	7	1.0	121	2	H82792	conserved hypothet
c	339	7	1.0	124	2	T04505	hypothetical prote
	340	7	1.0	125	2	E90762	probable regulator
c	341	7	1.0	126	2	S40341	Ig kappa chain - h
c	342	7	1.0	126	2	T00610	hypothetical prote
c	343	7	1.0	127	2	F72462	hypothetical prote
	344	7	1.0	130	2	T34729	probable gas vesic
	345	7	1.0	132	2	A86937	probable dihydrone
	346	7	1.0	132	2	F85625	hypothetical prote
c	347	7	1.0	133	2	F75552	ribosomal protein
c	348	7	1.0	133	2	S45752	hypothetical prote
	349	7	1.0	136	2	E84405	hypothetical prote
c	350	7	1.0	139	2	F97438	hypothetical prote
c	351	7	1.0	140	2	D90039	hypothetical prote
c	352	7	1.0	143	2	T34647	probable integral
c	353	7	1.0	143	2	T15722	hypothetical prote
c	354	7	1.0	144	2	B82623	TolR protein XF189
	355	7	1.0	144	2	F87452	transcription regu
	356	7	1.0	146	2	D95244	transcription regu
c	357	7	1.0	147	2	A53180	ribonuclease PL3 (
c	358	7	1.0	147	2	I52489	ribonuclease 4 (EC
c	359	7	1.0	149	2	S20316	synaptophysin homo
	360	7	1.0	150	2	F84976	3-dehydroquinate d
	361	7	1.0	150	2	D48835	kinesin-like prote
c	362	7	1.0	150	2	C86224	hypothetical prote
	363	7	1.0	153	2	AH2203	bacterioferritin c
c	364	7	1.0	153	2	B72721	hypothetical prote
c	365	7	1.0	155	2	B70370	hypothetical prote
	366	7	1.0	155	2	B97207	probable acetyltra
	367	7	1.0	156	2	D70827	ribosomal protein
	368	7	1.0	156	2	S31149	ribosomal protein
	369	7	1.0	156	2	A87144	30S ribosomal prot
	370	7	1.0	157	1	R3CWS7	ribosomal protein
	371	7	1.0	157	2	B26956	ribosomal protein
	372	7	1.0	157	2	H72065	ribosomal protein
	373	7	1.0	157	2	H81672	ribosomal protein
	374	7	1.0	157	2	S25479	homeotic protein H
	375	7	1.0	157	2	B84698	hypothetical prote
c	376	7	1.0	159	2	T05279	transcription fact
c	377	7	1.0	159	2	S37766	hypothetical prote
c	378	7	1.0	159	2	C75446	MutT/nudix family
c	379	7	1.0	159	2	S61040	probable membrane

c	380	7	1.0	161	2	AC0439	conserved hypothet
c	381	7	1.0	161	2	AG1132	hypothetical prote
	382	7	1.0	162	2	A45362	corticoliberin pre
	383	7	1.0	162	2	T33127	hypothetical prote
	384	7	1.0	165	1	RMEC18	primosomal operon
	385	7	1.0	165	2	S01510	NADH2 dehydrogenas
	386	7	1.0	165	2	H91293	probable glycoprot
	387	7	1.0	165	2	C86135	probable glycoprot
	388	7	1.0	166	2	C86559	S7 ribosomal prote
	389	7	1.0	167	2	T01155	hypothetical prote
c	390	7	1.0	167	2	A82641	hypothetical prote
c	391	7	1.0	168	2	C75425	hypothetical prote
	392	7	1.0	169	2	T26271	hypothetical prote
c	393	7	1.0	171	1	LWYC1	H+-transporting tw
	394	7	1.0	172	2	B98109	hypothetical prote
c	395	7	1.0	174	2	F83514	conserved hypothet
	396	7	1.0	174	2	T29067	probable ATP/GTP b
	397	7	1.0	174	2	T29071	probable ATP/GTP b
c	398	7	1.0	175	2	AC0853	hypothetical prote
	399	7	1.0	177	2	T00789	ubiquitin-protein
	400	7	1.0	177	2	H90631	probable fimbrial
	401	7	1.0	177	2	G85482	probable type-1 fi
	402	7	1.0	178	1	R3IT7	ribosomal protein
	403	7	1.0	178	2	A82606	conjugal transfer
	404	7	1.0	181	2	T51352	RNA helicase RH27
	405	7	1.0	181	2	T35858	probable ATP/GTP-b
	406	7	1.0	182	2	A46436	myofibrillar prote
c	407	7	1.0	183	2	PN0109	keratin-like prote
	408	7	1.0	184	2	A42749	hypothetical prote
	409	7	1.0	185	2	S39317	replication initia
	410	7	1.0	185	2	S39318	replication initia
	411	7	1.0	185	2	I40478	conserved hypothet
	412	7	1.0	187	2	B69536	acetylpolyamine am
c	413	7	1.0	188	2	C84787	hypothetical prote
	414	7	1.0	188	2	F84629	hypothetical prote
	415	7	1.0	189	2	H70416	hypothetical prote
c	416	7	1.0	190	2	S68230	antimicrobial pept
	417	7	1.0	191	2	T35637	probable ATP/GTP-b
c	418	7	1.0	192	2	JN0577	p-aminobenzoic aci
c	419	7	1.0	192	2	A70331	peptidyl-tRNA hydr
	420	7	1.0	194	2	T47200	H+-transporting tw
c	421	7	1.0	194	2	T10939	auxin-induced prot
c	422	7	1.0	196	2	S72716	4-coumarate-CoA li
c	423	7	1.0	202	2	D86034	probable S-transfe
c	424	7	1.0	202	2	E91187	probable S-transfe
c	425	7	1.0	202	2	S47813	hypothetical prote
	426	7	1.0	204	2	D91120	probable oxidoredu
	427	7	1.0	204	2	D85965	probable oxidoredu
c	428	7	1.0	205	2	S37804	hypothetical prote
	429	7	1.0	205	2	D72781	hypothetical prote
c	430	7	1.0	206	2	A53412	probable plastoqui
c	431	7	1.0	206	2	S18250	collagen alpha 1(I
	432	7	1.0	207	2	JH0145	hypothetical 24.1K
	433	7	1.0	209	2	B97543	hypothetical prote
	434	7	1.0	209	2	AC2762	conserved hypothet
	435	7	1.0	209	2	H65092	hypothetical prote
	436	7	1.0	210	2	S57943	cutinase - Phytoph
	437	7	1.0	210	2	S64376	hypothetical prote
c	438	7	1.0	214	2	T32066	hypothetical prote
	439	7	1.0	215	2	C69445	conserved hypothet
c	440	7	1.0	216	2	A71445	hypothetical prote
c	441	7	1.0	216	2	T45430	hypothetical prote
c	442	7	1.0	216	2	T46629	lp6 protein - lobl
	443	7	1.0	217	2	T18603	hypothetical prote
c	444	7	1.0	218	2	T28689	hypothetical prote
	445	7	1.0	221	2	A36370	immediate-early pr
c	446	7	1.0	221	2	T04592	glycine-rich cell
c	447	7	1.0	222	2	T43500	hypothetical prote
c	448	7	1.0	223	2	C70524	hypothetical prote
c	449	7	1.0	223	2	AB1110	hypothetical prote
	450	7	1.0	224	2	T33691	hypothetical prote
c	451	7	1.0	225	2	T44720	hypothetical prote
	452	7	1.0	228	1	C46335	vif protein - Maed

	453	7	1.0	228	2	AE0325	urease accessory p
	454	7	1.0	229	1	C45345	vif protein - capr
	455	7	1.0	229	2	JQ1932	vif protein - ovin
	456	7	1.0	230	1	QQLJVS	vif protein - Maed
	457	7	1.0	231	2	JQ1163	Vif protein - Maed
c	458	7	1.0	232	2	H87113	conserved hypothet
c	459	7	1.0	232	2	C85585	unknown protein en
c	460	7	1.0	232	2	B90735	hypothetical prote
	461	7	1.0	232	2	JQ0886	hypothetical prote
	462	7	1.0	233	2	B90088	similar to proteas
c	463	7	1.0	234	2	AE0505	fimbrial chaperone
c	464	7	1.0	234	2	E72252	translation initia
c	465	7	1.0	234	2	T49448	hypothetical prote
c	466	7	1.0	234	2	G95989	hypothetical glyci
	467	7	1.0	235	2	I40230	hypothetical prote
c	468	7	1.0	237	2	S55009	H+-transporting tw
	469	7	1.0	237	2	A98357	hypothetical prote
	470	7	1.0	237	2	AD2925	hypothetical prote
	471	7	1.0	238	2	T07955	H+-transporting tw
	472	7	1.0	239	2	C82870	hypothetical prote
c	473	7	1.0	239	2	A97358	glucose-inhibited
	474	7	1.0	240	2	E69004	hypothetical prote
c	475	7	1.0	240	2	AD1552	transcription regu
c	476	7	1.0	240	2	AF1194	transcription regu
	477	7	1.0	241	2	AH3543	gtp-binding protei
	478	7	1.0	242	2	C82575	phage-related prot
	479	7	1.0	243	2	JH0791	hypothetical prote
c	480	7	1.0	244	2	S18948	centrocyclin precu
	481	7	1.0	245	2	G72255	ABC transporter, A
c	482	7	1.0	245	2	JC5346	cdd2 protein - Clo
c	483	7	1.0	247	1	KYHUCM	chymase (EC 3.4.21
	484	7	1.0	247	2	T07213	H+-transporting tw
c	485	7	1.0	247	2	S58394	myelin/oligodendro
	486	7	1.0	247	2	H81719	conserved hypothet
	487	7	1.0	247	2	T30008	hypothetical prote
c	488	7	1.0	248	2	S33755	granzyme-like prot
c	489	7	1.0	249	2	B87309	hypothetical prote
	490	7	1.0	250	2	A49910	etrA protein - She
c	491	7	1.0	250	2	AC1171	hypothetical prote
	492	7	1.0	250	2	H83213	hypothetical prote
c	493	7	1.0	251	2	S40995	hypothetical prote
c	494	7	1.0	251	2	G70840	hypothetical prote
c	495	7	1.0	253	1	A40582	type IV prepilin p
	496	7	1.0	253	2	AB1717	dehydrogenase homo
	497	7	1.0	253	2	AG1346	dehydrogenase homo
c	498	7	1.0	253	2	D82273	leader peptidase T
c	499	7	1.0	253	2	T02106	hypothetical prote
c	500	7	1.0	254	2	AH3218	hypothetical prote
	501	7	1.0	254	2	T41954	hypothetical prote
c	502	7	1.0	256	2	S50280	PMP31 protein - ye
c	503	7	1.0	256	2	S50281	PMP32 protein - ye
	504	7	1.0	256	2	A90131	hypothetical prote
	505	7	1.0	257	2	F87209	probable enoyl-CoA
c	506	7	1.0	257	2	C84890	hypothetical prote
	507	7	1.0	258	1	A64451	probable 3-isoprop
c	508	7	1.0	259	2	I53171	pantophysin - huma
c	509	7	1.0	259	2	A71537	probable n-acetylm
c	510	7	1.0	259	2	F86475	hypothetical prote
	511	7	1.0	260	2	B71234	probable transport
	512	7	1.0	260	2	G75200	cobalt abc transpo
	513	7	1.0	262	2	B81865	probable fimbrial
c	514	7	1.0	262	2	T34691	hypothetical prote
	515	7	1.0	262	2	C88325	protein F43G6.5 [i
c	516	7	1.0	263	2	D75593	conserved hypothet
c	517	7	1.0	264	2	T05450	hypothetical prote
c	518	7	1.0	266	2	S51833	arcelin-4 precurso
c	519	7	1.0	267	2	A57490	matrilysin (EC 3.4
c	520	7	1.0	268	2	S09860	hypothetical prote
c	521	7	1.0	269	2	C71228	hypothetical prote
c	522	7	1.0	269	2	G86321	F6A14.11 protein -
c	523	7	1.0	270	2	A60830	keratin, 70k type
	524	7	1.0	270	2	D70171	ABC transporter, A
c	525	7	1.0	270	2	E98296	hypothetical prote

c	526	7	1.0	270	2	AC2987	conserved hypothet
c	527	7	1.0	273	2	T08645	yfjQ protein - Esc
c	528	7	1.0	273	2	F91083	hypothetical membr
c	529	7	1.0	273	2	H71212	hypothetical prote
c	530	7	1.0	274	2	C47700	glycerol uptake fa
	531	7	1.0	274	2	H98174	transcription regu
	532	7	1.0	274	2	AE3112	transcription regu
c	533	7	1.0	275	2	AC0069	cell division prot
c	534	7	1.0	275	2	D83626	transcription regu
c	535	7	1.0	276	2	F95259	conserved hypothet
c	536	7	1.0	276	2	A99525	conserved hypothet
	537	7	1.0	277	2	D86205	protein F10K1.23 [
	538	7	1.0	278	2	H86923	probable enoyl-CoA
c	539	7	1.0	278	2	D64750	yafZ protein - Esc
	540	7	1.0	278	2	T24430	hypothetical prote
c	541	7	1.0	279	2	T03830	probable myb facto
c	542	7	1.0	280	2	H71320	hypothetical prote
	543	7	1.0	281	2	T22670	hypothetical prote
c	544	7	1.0	281	2	T49537	hypothetical prote
	545	7	1.0	282	2	H87251	siroheme synthase-
c	546	7	1.0	286	1	EEWTA	alpha/beta-gliadin
c	547	7	1.0	286	2	S07923	alpha/beta-gliadin
c	548	7	1.0	286	2	F43670	integral membrane
c	549	7	1.0	288	2	T06282	alpha-gliadin prec
c	550	7	1.0	290	2	T31470	hypothetical prote
c	551	7	1.0	290	2	T31201	hypothetical prote
c	552	7	1.0	290	2	T42526	hypothetical prote
	553	7	1.0	291	2	AB1163	flagellar hook-ass
	554	7	1.0	291	2	AB1522	flagellar hook-ass
c	555	7	1.0	292	2	B22364	alpha/beta-gliadin
c	556	7	1.0	292	2	C71448	hypothetical prote
	557	7	1.0	293	2	H86718	primosomal protein
c	558	7	1.0	294	2	S23545	chitinase (EC 3.2.
	559	7	1.0	294	2	A97937	conserved hypothet
	560	7	1.0	294	2	A95069	leucine-rich prote
	561	7	1.0	295	2	F59102	probable UTP-gluco
	562	7	1.0	295	2	T51350	RNA helicase RH31
c	563	7	1.0	296	2	A27319	gliadin - wheat
c	564	7	1.0	296	2	S07361	alpha/beta-gliadin
c	565	7	1.0	296	2	S39587	peptide transport
c	566	7	1.0	296	2	G64877	peptide transport
c	567	7	1.0	296	2	E90862	hypothetical prote
c	568	7	1.0	296	2	D85756	hypothetical prote
c	569	7	1.0	296	2	AF0287	peptide transport
c	570	7	1.0	296	2	AI0656	peptide transport
c	571	7	1.0	296	2	T39265	hypothetical prote
c	572	7	1.0	297	2	T06500	alpha/beta-gliadin
	573	7	1.0	297	2	H83888	transmembrane lipo
	574	7	1.0	298	2	D95199	primosomal protein
	575	7	1.0	298	2	B98066	primosome componen
	576	7	1.0	298	2	T51746	RNA helicase RH17
c	577	7	1.0	299	2	T00837	glycine-rich prote
c	578	7	1.0	299	2	E82116	flagellar biosynth
c	579	7	1.0	299	2	T47989	RAV-like protein -
c	580	7	1.0	300	2	F90274	hypothetical prote
	581	7	1.0	300	2	T08453	hypothetical prote
c	582	7	1.0	301	2	H97505	probable nikC prot
c	583	7	1.0	301	2	AD2724	ABC transporter, m
	584	7	1.0	301	2	S72642	probable ABC-type
c	585	7	1.0	301	2	G85928	hypothetical prote
c	586	7	1.0	301	2	S35526	homeotic protein H
c	587	7	1.0	303	2	F89849	conserved hypothet
	588	7	1.0	303	2	F90301	hypothetical prote
	589	7	1.0	304	2	A45989	peroxisome assembl
	590	7	1.0	305	2	I52362	peroxisome assembl
	591	7	1.0	305	2	A41812	peroxisome assembl
	592	7	1.0	305	2	AC1588	hypothetical prote
	593	7	1.0	305	2	AH1651	hypothetical prote
c	594	7	1.0	305	2	H96832	hypothetical prote
	595	7	1.0	306	2	A54862	HSV-1 proteinase -
	596	7	1.0	306	2	A72610	hypothetical prote
c	597	7	1.0	307	2	S10015	alpha/beta-gliadin
	598	7	1.0	307	2	T03502	conserved hypothet
c	599	7	1.0	307	2	T39986	hypothetical prote
	600	7	1.0	308	2	A83914	ABC transporter (p
	601	7	1.0	309	2	T15696	hypothetical prote
	602	7	1.0	310	2	H81161	cysteine synthase
c	603	7	1.0	310	2	T02200	probable RAV-like
c	604	7	1.0	313	2	S07924	alpha/beta-gliadin
	605	7	1.0	313	2	JE0328	intelectin - mouse
c	606	7	1.0	313	2	H71438	hypothetical prote
	607	7	1.0	314	2	C81735	tRNA delta-2-isope
c	608	7	1.0	314	2	G97379	phoE protein U5922
c	609	7	1.0	315	2	D48560	immediate-early pr
c	610	7	1.0	316	1	A38743	loricrin - human
c	611	7	1.0	316	2	T41365	hypothetical prote
c	612	7	1.0	316	2	E96770	hypothetical prote
c	613	7	1.0	319	2	A22364	alpha/beta-gliadin
c	614	7	1.0	319	2	T01848	hypothetical prote
c	615	7	1.0	320	2	T35265	probable D-amino a
	616	7	1.0	320	2	C96922	probable acetyltra
	617	7	1.0	321	2	T34065	hypothetical prote
	618	7	1.0	321	2	C90402	transposase ISC117
	619	7	1.0	321	2	D90326	transposase ISC117
	620	7	1.0	324	2	B69521	hypothetical prote
c	621	7	1.0	324	2	I50125	cdx1 protein - com
c	622	7	1.0	325	2	T14227	NADH2 dehydrogenas
c	623	7	1.0	325	2	AF2665	conserved hypothet
c	624	7	1.0	326	2	D22364	alpha/beta-gliadin
	625	7	1.0	326	2	E71047	probable signal re
c	626	7	1.0	326	2	C90272	hypothetical prote
c	627	7	1.0	327	2	G70789	probable epoxide h
c	628	7	1.0	328	2	G88499	protein K04G7.10 [
	629	7	1.0	328	2	T36574	probable partition
c	630	7	1.0	328	2	H97089	uncharacterized AT
c	631	7	1.0	328	2	AF2597	hypothetical prote
c	632	7	1.0	329	2	E69456	hypothetical prote
	633	7	1.0	330	2	B48835	kinesin-like prote
c	634	7	1.0	331	2	C95998	probable sugar upt
	635	7	1.0	331	2	G75099	sugar transferase
	636	7	1.0	332	2	T01483	hypothetical prote
c	637	7	1.0	333	2	E97447	hypothetical prote
c	638	7	1.0	333	2	T33649	hypothetical prote
	639	7	1.0	333	2	F83107	hypothetical prote
c	640	7	1.0	334	2	B72033	conserved hypothet
c	641	7	1.0	334	2	C86593	CT566 hypothetical
c	642	7	1.0	334	2	T36485	probable araC-fami
c	643	7	1.0	335	2	F72404	flagellar motor sw
c	644	7	1.0	336	2	S35156	beta-glucanase - b
c	645	7	1.0	336	2	S72858	hypothetical prote
c	646	7	1.0	337	2	I47079	follistatin - shee
	647	7	1.0	339	2	F47301	VirB11 homolog - B
	648	7	1.0	339	2	A71473	probable tRNA pyro
	649	7	1.0	339	2	T25315	hypothetical prote
c	650	7	1.0	343	2	S55369	follistatin - chic
c	651	7	1.0	343	2	S45321	follistatin - mous
	652	7	1.0	343	2	T33945	hypothetical prote
c	653	7	1.0	344	1	A27701	follistatin precur
	654	7	1.0	344	2	T12337	NADH2 dehydrogenas
c	655	7	1.0	344	2	A32141	follistatin 1 prec
c	656	7	1.0	344	2	I45894	follistatin - bovi
c	657	7	1.0	344	2	I57698	follistatin - rat
	658	7	1.0	345	2	T12339	NADH2 dehydrogenas
c	659	7	1.0	345	2	D88485	protein F23F12.4 [
c	660	7	1.0	346	2	AI2229	cobalamin synthesi
c	661	7	1.0	348	2	T07660	probable serine/th
c	662	7	1.0	348	2	G02297	gene N33 protein -
c	663	7	1.0	348	2	A83250	O-antigen chain le
	664	7	1.0	348	2	S34494	ccsA protein - Eug
	665	7	1.0	348	2	T17304	hypothetical prote
c	666	7	1.0	349	2	F84246	hypothetical prote
	667	7	1.0	350	2	E87714	hypothetical prote
	668	7	1.0	351	2	C75479	conserved hypothet
c	669	7	1.0	351	2	G83799	two-component sens
c	670	7	1.0	352	2	T06747	hypothetical prote
	671	7	1.0	352	2	T35891	hypothetical prote

	672	7	1.0	352	2	T33664	hypothetical prote
c	673	7	1.0	353	2	AD3169	transcription regu
	674	7	1.0	353	2	H75020	hypothetical prote
c	675	7	1.0	355	2	T04317	homeobox protein L
c	676	7	1.0	355	2	S06939	hypothetical prote
c	677	7	1.0	356	2	T48206	protein kinase ATN
c	678	7	1.0	356	2	S39889	SMR1 protein - Pod
	679	7	1.0	356	2	H90168	GTP-binding protei
c	680	7	1.0	358	2	T26281	hypothetical prote
c	681	7	1.0	361	2	A48373	high-alkaline seri
c	682	7	1.0	361	2	G83756	subtilisin-type al
	683	7	1.0	361	2	F91207	hypothetical prote
	684	7	1.0	361	2	C72751	hypothetical prote
	685	7	1.0	361	2	H86053	hypothetical prote
c	686	7	1.0	363	2	C86214	hypothetical prote
	687	7	1.0	363	2	H97575	probable permease
	688	7	1.0	363	2	AH2796	hypothetical prote
	689	7	1.0	363	2	H81334	hypothetical prote
	690	7	1.0	363	2	T51341	RNA helicase RH9 [
c	691	7	1.0	363	2	H97176	lysophospholipase
c	692	7	1.0	364	2	AH1857	anthranilate phosp
	693	7	1.0	365	2	H96673	hypothetical prote
c	694	7	1.0	366	1	JC1106	inhibin alpha chai
c	695	7	1.0	366	1	A40056	inhibin alpha chai
	696	7	1.0	366	2	F90474	hypothetical prote
c	697	7	1.0	367	2	E84453	probable GDSL-moti
	698	7	1.0	368	2	G82321	twitching motility
	699	7	1.0	369	2	A39157	probable RNA helic
	700	7	1.0	369	2	S76435	hypothetical prote
	701	7	1.0	369	2	B70968	hypothetical prote
c	702	7	1.0	371	2	T13021	hypothetical prote
c	703	7	1.0	371	2	H70326	hypothetical prote
	704	7	1.0	372	2	T10202	hypothetical prote
c	705	7	1.0	373	1	DERTH1	3beta-hydroxy-Delt
c	706	7	1.0	373	1	DERTH2	3beta-hydroxy-Delt
c	707	7	1.0	373	1	DERTHM	3beta-hydroxy-Delt
c	708	7	1.0	373	2	A48769	3beta-hydroxystero
c	709	7	1.0	373	2	A37404	3beta-hydroxy-Delt
c	710	7	1.0	373	2	A49573	3beta-hydroxy-Delt
c	711	7	1.0	373	2	I49762	3beta-hydroxy-Delt
c	712	7	1.0	373	2	JC7289	G-protein coupled
c	713	7	1.0	374	2	E95361	probable muconate
c	714	7	1.0	375	2	AC0959	hypothetical prote
c	715	7	1.0	376	2	D84577	probable recA prot
c	716	7	1.0	378	2	T00481	probable RING zinc
c	717	7	1.0	379	1	QQEC4W	Citrate utilizatio
c	718	7	1.0	379	2	AF2409	mannosyl transfera
c	719	7	1.0	379	2	G75170	hypothetical prote
	720	7	1.0	380	1	NPBY	anthranilate phosp
c	721	7	1.0	380	2	I50701	transcription fact
c	722	7	1.0	381	2	G96804	hypothetical prote
c	723	7	1.0	381	2	B86175	protein F19P19.21
	724	7	1.0	381	2	AI0622	hypothetical proph
	725	7	1.0	382	2	T27058	hypothetical prote
c	726	7	1.0	383	2	A42111	Na+/H+-exchanging
	727	7	1.0	386	2	F90112	DEAD box protein [
	728	7	1.0	387	2	T48731	probable translati
	729	7	1.0	389	2	D84068	RNA helicase BH334
c	730	7	1.0	389	2	A82327	MSHA biogenesis pr
	731	7	1.0	390	2	G01936	Abl binding protei
	732	7	1.0	391	2	G82320	probable oxygen-in
	733	7	1.0	393	2	G96781	unknown protein F2
c	734	7	1.0	393	2	JC5275	voltage-gated pota
c	735	7	1.0	393	2	JC5614	RNB6 protein - rat
c	736	7	1.0	395	2	AC3571	acyl-CoA dehydroge
	737	7	1.0	396	2	I38430	connexin45 - human
c	738	7	1.0	396	2	T18854	hypothetical prote
	739	7	1.0	397	1	JC4017	triacylglycerol li
c	740	7	1.0	397	2	B98293	mannonate dehydrat
c	741	7	1.0	397	2	AG2990	mannonate dehydrat
c	742	7	1.0	397	2	T17008	knotted1-like home
c	743	7	1.0	397	2	T13516	hypothetical prote
c	744	7	1.0	398	2	A81717	conserved hypothet
c	745	7	1.0	398	2	T17009	knotted1-like home
	746	7	1.0	399	2	T32773	hypothetical prote
	747	7	1.0	399	2	S54644	translation initia
c	748	7	1.0	399	2	G75024	hypothetical prote
c	749	7	1.0	399	2	B71140	hypothetical prote
c	750	7	1.0	399	2	T48267	probable zinc fing
c	751	7	1.0	401	2	AI3568	mannonate dehydrat
c	752	7	1.0	403	2	A26371	sex steroid-bindin
	753	7	1.0	403	2	A81464	translation initia
c	754	7	1.0	404	2	S54729	RNA-binding protei
c	755	7	1.0	404	2	G70886	probable integral
	756	7	1.0	405	2	D84950	acetate kinase (EC
	757	7	1.0	406	2	G97185	membrane protein o
	758	7	1.0	407	2	AC2111	pilin biogenesis p
	759	7	1.0	407	2	D82377	ParA family protei
c	760	7	1.0	410	1	S68153	cellulase (EC 3.2.
c	761	7	1.0	410	2	S18157	globulin 2 - easte
	762	7	1.0	411	2	S45142	translation initia
	763	7	1.0	412	2	JC1453	translation initia
	764	7	1.0	412	2	JC1452	translation initia
c	765	7	1.0	412	2	S62538	hypothetical coile
c	766	7	1.0	413	2	T47272	transposase tnpA [
	767	7	1.0	413	2	T24767	hypothetical prote
c	768	7	1.0	413	2	AB3174	aminotransferase,
	769	7	1.0	414	2	B96752	hypothetical prote
c	770	7	1.0	414	2	T38742	hypothetical prote
c	771	7	1.0	414	2	A40350	transcription repr
c	772	7	1.0	415	2	E83377	probable alcohol d
	773	7	1.0	416	2	F83197	probable porin PA3
c	774	7	1.0	416	2	S50342	homeotic protein S
c	775	7	1.0	416	2	G83656	hypothetical prote
c	776	7	1.0	419	2	T43572	type III secretion
c	777	7	1.0	420	2	I51667	thrombin receptor
	778	7	1.0	423	2	A90450	conserved hypothet
c	779	7	1.0	423	2	T03269	uroporphyrin-III C
	780	7	1.0	425	2	S17759	protein kinase, ca
	781	7	1.0	427	2	B64909	probable membrane
	782	7	1.0	427	2	H90897	probable transport
	783	7	1.0	427	2	G85719	probable transport
c	784	7	1.0	427	2	A32372	female-specific do
c	785	7	1.0	431	2	D86318	protein F15H18.6 [
	786	7	1.0	431	2	F85097	hypothetical prote
c	787	7	1.0	431	2	T04868	hypothetical prote
c	788	7	1.0	432	2	T39816	hypothetical prote
	789	7	1.0	432	2	D87649	tolB protein [impo
c	790	7	1.0	432	2	T23561	hypothetical prote
c	791	7	1.0	433	1	QQBEV2	UL44 protein - hum
c	792	7	1.0	433	2	T11162	NADH2 dehydrogenas
	793	7	1.0	433	2	S07586	phosphopyruvate hy
c	794	7	1.0	433	2	S20963	homeotic protein H
	795	7	1.0	434	2	S10246	phosphopyruvate hy
	796	7	1.0	434	2	A23126	phosphopyruvate hy
c	797	7	1.0	434	2	S62168	hypothetical prote
	798	7	1.0	435	2	AB1786	cellobiose phospho
	799	7	1.0	435	2	AC1410	cellobiose phospho
c	800	7	1.0	435	2	S75346	hypothetical prote
c	801	7	1.0	435	2	T08454	hypothetical prote
c	802	7	1.0	436	2	JC5599	cholecystokinin-A
c	803	7	1.0	437	2	S59151	NADH2 dehydrogenas
c	804	7	1.0	438	2	AD0389	phosphate regulon
c	805	7	1.0	438	2	T07790	transaldolase (EC
	806	7	1.0	438	2	B81410	probable integral
c	807	7	1.0	439	2	E84780	hypothetical prote
c	808	7	1.0	440	2	T11358	NADH2 dehydrogenas
	809	7	1.0	441	2	T51347	RNA helicase RH23
	810	7	1.0	441	2	D71191	hypothetical prote
	811	7	1.0	441	2	F75150	tbp-interacting pr
	812	7	1.0	442	2	A71433	hypothetical prote
	813	7	1.0	442	2	H98243	agaE protein [impo
	814	7	1.0	442	2	AD3042	oxidoreductase Atu
c	815	7	1.0	443	2	S37612	NADH2 dehydrogenas
c	816	7	1.0	444	2	T42674	hypothetical prote
	817	7	1.0	445	2	S39203	phosphopyruvate hy

c	818	7	1.0	445	2	A57400	embryonic TEA doma
c	819	7	1.0	445	2	T05639	hypothetical prote
c	820	7	1.0	446	2	T45525	WSC4 homolog [impo
c	821	7	1.0	447	2	AF2145	hypothetical prote
	822	7	1.0	448	2	E89936	hypothetical prote
c	823	7	1.0	449	1	H70526	probable cytochrom
	824	7	1.0	449	2	I64246	ATP-dependent RNA
	825	7	1.0	450	2	S73545	ATP-dependent RNA
	826	7	1.0	450	2	H87493	conserved hypothet
	827	7	1.0	451	2	S56717	calcium-dependent
	828	7	1.0	451	2	G82096	sodium-dependent t
c	829	7	1.0	452	2	C71391	NADH2 dehydrogenas
c	830	7	1.0	452	2	B87075	probable DNA-bindi
c	831	7	1.0	452	2	D70768	hypothetical prote
c	832	7	1.0	456	2	B90543	hypothetical prote
c	833	7	1.0	459	2	T11110	NADH2 dehydrogenas
c	834	7	1.0	459	2	D96833	hypothetical prote
	835	7	1.0	459	2	AE0679	probable membrane
	836	7	1.0	459	2	B95009	potassium uptake p
	837	7	1.0	459	2	F97880	hypothetical prote
c	838	7	1.0	460	2	T13881	NADH2 dehydrogenas
c	839	7	1.0	461	2	H84099	cell wall-binding
	840	7	1.0	463	2	T51194	hypothetical prote
c	841	7	1.0	464	2	JH0824	gamma-aminobutyric
	842	7	1.0	464	2	S22697	extensin - Volvox
	843	7	1.0	465	1	T03024	calcium-dependent
c	844	7	1.0	465	1	WZBEB4	51.3K capsid prote
c	845	7	1.0	465	2	AC0347	probable membrane
c	846	7	1.0	465	2	T34984	probable integral
c	847	7	1.0	465	2	T19113	hypothetical prote
	848	7	1.0	465	2	T27415	hypothetical prote
c	849	7	1.0	466	2	C42360	cellulase (EC 3.2.
c	850	7	1.0	467	2	S38535	cytochrome P450 76
c	851	7	1.0	469	1	I38950	DNA-directed DNA p
	852	7	1.0	469	2	E82264	probable polysacch
c	853	7	1.0	471	2	T20690	hypothetical prote
c	854	7	1.0	473	2	E72723	probable TATA-bind
c	855	7	1.0	473	2	G82405	transcription regu
c	856	7	1.0	474	1	T05239	acid phosphatase (
c	857	7	1.0	474	2	G84372	dihydrolipoamide d
c	858	7	1.0	474	2	S63658	NADH2 dehydrogenas
	859	7	1.0	474	2	T39587	serine/threonine-p
c	860	7	1.0	475	1	A56824	dihydrolipoamide d
c	861	7	1.0	476	2	C96577	hypothetical prote
c	862	7	1.0	476	2	H84524	probable fatty aci
c	863	7	1.0	476	2	T32032	hypothetical prote
	864	7	1.0	478	2	G75530	probable ATP-depen
c	865	7	1.0	480	2	F86207	hypothetical prote
c	866	7	1.0	480	2	A31589	carboxypeptidase C
c	867	7	1.0	481	2	JE0377	p70 S6 kinase (EC
c	868	7	1.0	481	2	A35628	loricrin - mouse
	869	7	1.0	481	2	C71079	hypothetical prote
c	870	7	1.0	482	2	S60757	catalase (EC 1.11.
	871	7	1.0	482	2	S27608	cysteine proteinas
c	872	7	1.0	483	2	H87492	YjeF family protei
	873	7	1.0	484	2	T05650	calcium-dependent
c	874	7	1.0	484	2	F88924	protein R02C2.2 [i
c	875	7	1.0	485	2	JQ1957	glucagon receptor
	876	7	1.0	485	2	AG2531	hypothetical prote
	877	7	1.0	487	1	S71770	calcium-dependent
	878	7	1.0	487	2	E69895	xylulokinase homol
	879	7	1.0	487	2	AB1204	glycerol kinase ho
c	880	7	1.0	488	2	S18156	globulin 1 - easte
	881	7	1.0	489	1	NIBCAT	nitrogenase (EC 1.
c	882	7	1.0	489	2	AG3038	conserved hypothet
c	883	7	1.0	490	1	S71776	calcium-dependent
	884	7	1.0	490	2	D70008	nicotinate phospho
	885	7	1.0	490	2	JC5641	sugar transporter
	886	7	1.0	490	2	F89824	hypothetical prote
c	887	7	1.0	491	2	JC6197	stromelysin 3 (EC
c	888	7	1.0	491	2	T48991	hypothetical prote
	889	7	1.0	491	2	S63206	LET1 protein homol
	890	7	1.0	492	1	T03271	calcium-dependent
c	891	7	1.0	492	2	A44399	stromelysin 3 (EC
c	892	7	1.0	493	2	S78183	NADH2 dehydrogenas
c	893	7	1.0	493	2	G84263	long-chain fatty-a
c	894	7	1.0	494	2	AB1326	two-component resp
c	895	7	1.0	496	2	JC5110	cyclin-dependent k
c	896	7	1.0	497	1	A43349	steroid 21-monooxy
c	897	7	1.0	497	1	S53834	NADH2 dehydrogenas
	898	7	1.0	497	2	T14433	reproductive meris
c	899	7	1.0	497	2	S22708	homeotic protein e
c	900	7	1.0	498	2	T11039	NADH2 dehydrogenas
c	901	7	1.0	498	2	T13434	hypothetical prote
c	902	7	1.0	500	2	A53658	prostaglandin-I sy
c	903	7	1.0	500	2	AC1915	NADH dehydrogenase
c	904	7	1.0	502	2	AD0691	conserved hypothet
c	905	7	1.0	504	2	T07415	probable serine/th
	906	7	1.0	506	1	W2WL47	E2 protein - human
c	907	7	1.0	507	2	S75961	NADH2 dehydrogenas
c	908	7	1.0	507	2	T31975	hypothetical prote
	909	7	1.0	507	2	T24944	hypothetical prote
c	910	7	1.0	509	2	T39542	hypothetical prote
c	911	7	1.0	511	2	T11940	ribosomal protein
	912	7	1.0	512	2	T40576	probable structure
	913	7	1.0	513	1	T02259	calcium-dependent
c	914	7	1.0	513	2	T48788	hypothetical prote
	915	7	1.0	514	2	T10938	calcium-dependent
	916	7	1.0	514	2	T44976	hydantoinase homol
c	917	7	1.0	515	2	D98247	mexT protein (AF11
	918	7	1.0	515	2	AC2635	glycosyltransferas
c	919	7	1.0	516	2	AC1540	ATP-dependent RNA
	920	7	1.0	516	2	S40996	hypothetical prote
c	921	7	1.0	518	2	T11456	cytochrome-c oxida
	922	7	1.0	518	2	T25507	hypothetical prote
c	923	7	1.0	518	2	S46784	hypothetical prote
	924	7	1.0	519	2	I54523	kinesin-related pr
	925	7	1.0	520	2	I51556	recombination acti
c	926	7	1.0	520	2	AB1183	ATP-dependent RNA
c	927	7	1.0	521	1	HYBS	bacillolysin (EC 3
	928	7	1.0	521	2	G96543	calcium-dependent
c	929	7	1.0	521	2	JQ2129	bacillolysin (EC 3
c	930	7	1.0	521	2	T01242	peptide transport
c	931	7	1.0	522	2	D82195	phage replication
c	932	7	1.0	523	2	D90270	maltose ABC transp
c	933	7	1.0	525	2	T31802	hypothetical prote
	934	7	1.0	525	2	T48824	hypothetical prote
c	935	7	1.0	526	1	KRBOVI	keratin, 54K type
c	936	7	1.0	526	2	E65024	Hydrogenase-4 comp
c	937	7	1.0	526	2	D91047	hydrogenase 4 memb
c	938	7	1.0	526	2	H85891	hydrogenase 4 memb
c	939	7	1.0	526	2	C83790	hypothetical prote
c	940	7	1.0	527	2	S26037	NADH2 dehydrogenas
c	941	7	1.0	528	2	G02127	fus-like protein -
	942	7	1.0	528	2	S42510	Rag-2 protein - ch
c	943	7	1.0	529	2	S46116	probable regulator
c	944	7	1.0	530	2	T48627	hypothetical prote
	945	7	1.0	531	1	T02993	calcium-dependent
	946	7	1.0	531	2	D85059	probable calcium d
	947	7	1.0	531	2	B83082	probable binding p
	948	7	1.0	532	2	T14335	protein kinase, ca
	949	7	1.0	533	1	S56652	calcium-dependent
c	950	7	1.0	533	2	JS0304	developmental cont
	951	7	1.0	534	1	JC1515	calcium-dependent
c	952	7	1.0	534	2	H71069	probable nodulatio
	953	7	1.0	538	2	T08874	calcium-dependent
c	954	7	1.0	538	2	S67766	RNA-export mediato
c	955	7	1.0	539	2	B72707	hypothetical prote
	956	7	1.0	540	1	T01989	calcium-dependent
	957	7	1.0	541	2	S51799	nucleoporin NUP57
	958	7	1.0	542	1	S56651	calcium-dependent
c	959	7	1.0	542	2	AG0341	hypothetical prote
c	960	7	1.0	542	2	D83041	probable two-compo
	961	7	1.0	544	2	D84550	probable calmoduli
	962	7	1.0	544	2	S43058	CCTeta protein eta
c	963	7	1.0	545	2	T45935	probable galactose

	964	7	1.0	546	2	A84630	hypothetical prote
c	965	7	1.0	547	2	B97114	probable kinase re
	966	7	1.0	547	2	S44841	K06H7.1 protein -
	967	7	1.0	548	2	T47510	probable transport
c	968	7	1.0	549	2	B32372	male-specific doub
	969	7	1.0	550	2	G70597	probable proteinas
	970	7	1.0	551	2	S65289	hypothetical prote
	971	7	1.0	551	2	S66740	probable transcrip
	972	7	1.0	553	1	T02139	calcium-dependent
c	973	7	1.0	553	1	B55483	transcription init
	974	7	1.0	554	1	T03263	calcium-dependent
	975	7	1.0	554	2	T05476	calcium-dependent
	976	7	1.0	554	2	T22140	hypothetical prote
	977	7	1.0	554	2	E87375	conserved hypothet
	978	7	1.0	554	2	T50118	kinesin-related pr
	979	7	1.0	556	2	T06126	calcium-dependent
c	980	7	1.0	556	2	G90352	thermopsin precurs
c	981	7	1.0	556	2	T16790	hypothetical prote
	982	7	1.0	557	1	NUVKL	glucose-6-phosphat
c	983	7	1.0	559	2	T22179	hypothetical prote
c	984	7	1.0	559	2	JT0949	egg-specific prote
	985	7	1.0	560	2	F81423	L-lactate permease
	986	7	1.0	560	2	T08433	helicase homolog h
	987	7	1.0	562	2	C71604	RNA helicase PFB08
	988	7	1.0	563	2	T51348	RNA helicase RH25
c	989	7	1.0	564	2	C42523	A55R protein - vac
c	990	7	1.0	564	2	JQ1792	SalF17R protein -
	991	7	1.0	565	2	S75255	tyrS protein slr10
c	992	7	1.0	565	2	H86477	protein F15O4.19 [
c	993	7	1.0	568	2	G02753	testis specific ba
	994	7	1.0	570	2	C86440	PPR-repeat protein
	995	7	1.0	572	2	T16780	hypothetical prote
	996	7	1.0	572	2	H75271	conserved hypothet
c	997	7	1.0	575	2	T48224	probable homeodoma
	998	7	1.0	577	2	S77715	cytochrome c-type
c	999	7	1.0	578	2	A70877	probable acyl-coAs
	1000	7	1.0	578	2	AG0119	probable OmpA-fami

ALIGNMENTS

RESULT 1
T46407
probable RNA helicase protein DKFZp434B1122.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
C;Accession: T46407
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23034
A;Accession: T46407
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-635 <AAA>
A;Cross-references: EMBL:AL137462; PIDN:CAB70750.1
A;Experimental source: adult testis; clone DKFZp434B1122
C;Genetics:
A;Note: DKFZp434B1122.1

Alignment Scores:

Pred. No.:	0	Length:	635
Score:	522.00	Matches:	622
Percent Similarity:	99.84%	Conservative:	0
Best Local Similarity:	99.84%	Mismatches:	1
Query Match:	72.10%	Indels:	0
DB:	2	Gaps:	0

US-09-714-865-15 (1-2172) x T46407 (1-635)

```
Qy    268 AGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTC 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 SerPheGlyAsnArgGlyPheSerAsnSerArgPheGluAspGlyAspSerSerGlyPhe 20

Qy    328 TGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAG 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 TrpArgGluSerSerAsnAspCysGluAspAsnProThrArgAsnArgGlyPheSerLys 40

Qy    388 AGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGA 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 ArgGlyGlyTyrArgAspGlyAsnAsnSerGluAlaSerGlyProTyrArgArgGlyGly 60

Qy    448 AGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTA 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ArgGlySerPheArgGlyCysArgGlyGlyPheGlyLeuGlySerProAsnAsnAspLeu 80

Qy    508 GACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTA 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 AspProAspGluCysMetGlnArgThrGlyGlyLeuPheGlySerArgArgProValLeu 100

Qy    568 AGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGT 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 SerGlyThrGlyAsnGlyAspThrSerGlnSerArgSerGlySerGlySerGluArgGly 120

Qy    628 GGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCA 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlyTyrLysGlyLeuAsnGluGluValIleThrGlySerGlyLysAsnSerTrpLysSer 140

Qy    688 GAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCT 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GluAlaGluGlyGlyGluSerSerAspThrGlnGlyProLysValThrTyrIleProPro 160

Qy    748 CCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGAC 807
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 ProProProGluAspGluAspSerIlePheAlaHisTyrGlnThrGlyIleAsnPheAsp 180

Qy    808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysTyrAspThrIleLeuValGluValSerGlyHisAspAlaProProAlaIleLeuThr 200

Qy    868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 PheGluGluAlaAsnLeuCysGlnThrLeuAsnAsnAsnIleAlaLysAlaGlyTyrThr 220

Qy    928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 LysLeuThrProValGlnLysTyrSerIleProIleIleLeuAlaGlyArgAspLeuMet 240

Qy    988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuAlaHis 260

Qy   1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 MetMetHisAspGlyIleThrAlaSerArgPheLysGluLeuGlnGluProGluCysIle 280

Qy   1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 IleValAlaProThrArgGluLeuValAsnGlnIleTyrLeuGluAlaArgLysPheSer 300

Qy   1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PheGlyThrCysValArgAlaValValIleTyrGlyGlyThrGlnLeuGlyHisSerIle 320

Qy   1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 ArgGlnIleValGlnGlyCysAsnIleLeuCysAlaThrProGlyArgLeuMetAspIle 340

Qy   1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 IleGlyLysGluLysIleGlyLeuLysGlnIleLysTyrLeuValLeuAspGluAlaAsp 360

Qy   1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ArgMetLeuAspMetGlyPheGlyProGluMetLysLysLeuIleSerCysProGlyMet 380

Qy   1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ProSerLysGluGlnArgGlnThrLeuMetPheSerAlaThrPheProGluGluIleGln 400

Qy       1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGT 1527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 ArgLeuAlaAlaGluPheLeuLysSerAsnTyrLeuPheValAlaValGlyGlnValGly 420

Qy       1528 GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAA 1587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GlyAlaCysArgAspValGlnGlnThrValLeuGlnValGlyGlnPheSerLysArgGlu 440

Qy       1588 AAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACT 1647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LysLeuValGluIleLeuArgAsnIleGlyAspGluArgThrMetValPheValGluThr 460

Qy       1648 AAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGT 1707
              |||||||||||||||||      |||||||||||||||||||||||||||||||||||||
Db        461 LysLysLysAlaAspPheIleAlaThrPheLeuCysGlnGluLysIleSerThrThrSer 480

Qy       1708 ATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGA 1767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 IleHisGlyAspArgGluGlnArgGluArgGluGlnAlaLeuGlyAspPheArgPheGly 500

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LysCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluAsnVal 520

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GlnHisValIleAsnPheAspLeuProSerThrIleAspGluTyrValHisArgIleGly 540

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGluSerAsp 560

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 AsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValProAla 580

Qy       2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 TrpLeuGluGluIleAlaPheSerThrTyrIleProGlyPheSerGlySerThrArgGly 600

Qy       2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AsnValPheAlaSerValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGlyPhe 620

Qy       2128 TCTTCTTCA 2136
              |||||||||
Db        621 SerSerSer 623
```

RESULT 2
JC2534
RVLG protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-May-1995 #sequence_revision 14-Jul-1995 #text_change 02-Feb-2001
C;Accession: JC2534
R;Komiya, T.; Tanigawa, Y.
Biochem. Biophys. Res. Commun. 207, 405-410, 1995
A;Title: Cloning of a gene of the DEAD Box protein family which is specifically expresse
A;Reference number: JC2534; MUID:95160706; PMID:7857296
A;Accession: JC2534
A;Molecule type: mRNA
A;Residues: 1-713 <KOM>
A;Cross-references: GB:S75275; NID:g806463; PIDN:AAB33364.1; PID:g806464
C;Comment: This protein contains a conserved DEAD box.
C;Keywords: ATP; nucleotide binding; P-loop
F;317-324/Region: nucleotide-binding motif A (P-loop)
F;427-432/Region: nucleotide-binding motif B
F;431-434/Region: DEAD motif

```
Alignment Scores:
Pred. No.:              1.57e-74       Length:        713
Score:                  82.00          Matches:       82
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            11.33%         Indels:        0
DB:                     2              Gaps:          0

US-09-714-865-15 (1-2172) x JC2534 (1-713)

Qy        718 CAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePhe 244

Qy        778 GCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 AlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSer 264

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeu 284

Qy        898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 AsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIle 304

Qy        958 CCTATC 963
              ||||||
Db        305 ProIle 306
```

RESULT 3
I49638
probable RNA helicase protein - mouse (fragment)
N;Alternate names: Drosophila vasa homolog
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Feb-2001
C;Accession: I49638
R;Fujiwara, Y.; Komiya, T.; Kawabata, H.; Sato, M.; Fujimoto, H.; Furusawa, M.; Noce, T.
Proc. Natl. Acad. Sci. U.S.A. 91, 12258-12262, 1994
A;Title: Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosopl
A;Reference number: I49638; MUID:95083681; PMID:7991615
A;Accession: I49638
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-637 <RES>
A;Cross-references: GB:D14859; NID:g286074; PIDN:BAA03584.1; PID:g286075
C;Keywords: ATP; nucleotide binding; P-loop
F;246-253/Region: nucleotide-binding motif A (P-loop)
F;356-361/Region: nucleotide-binding motif B
F;360-363/Region: DEAD motif

```
Alignment Scores:
Pred. No.:              3.76e-56       Length:        637
Score:                  64.00          Matches:       64
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            8.84%          Indels:        0
DB:                     2              Gaps:          0

US-09-714-865-15 (1-2172) x I49638 (1-637)

Qy        718 CAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePhe 173

Qy        778 GCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSer 193

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeu 213

Qy        898 AATAACAACATT 909
              ||||||||||||
```

```
Db        214 AsnAsnAsnIle 217

RESULT 4
I51235
DEAD box protein - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Feb-2001
C;Accession: I51235
R;Komiya, T.; Itoh, K.; Ikenishi, K.; Furusawa, M.
Dev. Biol. 162, 354-363, 1994
A;Title: Isolation and characterization of a novel gene of the DEAD box protein family w
A;Reference number: I51235; MUID:94200507; PMID:8150200
A;Accession: I51235
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-700 <KOM>
A;Cross-references: GB:S69534; NID:g2896106; PIDN:AAC03114.1; PID:g2896107
C;Genetics:
A;Gene: XVLG1
C;Keywords: ATP; nucleotide binding; P-loop
F;318-325/Region: nucleotide-binding motif A (P-loop)
F;428-433/Region: nucleotide-binding motif B
F;432-435/Region: DEAD motif

Alignment Scores:
Pred. No.:              2.51e-15      Length:        700
Score:                  24.00         Matches:       24
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.31%         Indels:        0
DB:                     2             Gaps:          0

US-09-714-865-15 (1-2172) x I51235 (1-700)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 329

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        330 LeuProIleLeu 333

RESULT 5
A58768
ATP-dependent RNA helicase homolog - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 31-Dec-1990 #sequence_revision 17-Apr-1998 #text_change 19-Jan-2001
C;Accession: A58768; S01676; S10129; A31922
R;Liang, L.; Diehl-Jones, W.; Lasko, P.
unpublished results, 1995, cited by GenBank in release 103.0
A;Description: Localization of Vasa protein to the Drosophila pole plasm is independent
A;Reference number: A58768
A;Accession: A58768
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-661 <LAS1>
A;Cross-references: EMBL:X12945; NID:g433675; PIDN:CAA31405.1; PID:g1054723
R;Lasko, P.F.; Ashburner, M.
Nature 335, 611-617, 1988
A;Title: The product of the Drosophila gene vasa is very similar to eukaryotic initiatic
A;Reference number: S01676; MUID:89014721; PMID:3140040
A;Accession: S01676
A;Molecule type: DNA
A;Residues: 1-34,'R',36-448,'LRF',453-460,'R',462-589,'Q',591-661 <LAS2>
A;Cross-references: EMBL:X12945; NID:g433675
R;Ashburner, M.
submitted to the EMBL Data Library, October 1988
A;Reference number: S10129
A;Accession: S10129
A;Molecule type: DNA
A;Residues: 1-34,'R',36-644,'Q',645-661 <ASH>
A;Cross-references: EMBL:X12945; NID:g433675
R;Hay, B.; Jan, L.Y.; Jan, Y.N.
Cell 55, 577-587, 1988
A;Title: A protein component of Drosophila polar granules is encoded by vasa and has exte
A;Reference number: A31922; MUID:89028669; PMID:3052853
A;Accession: A31922
A;Molecule type: mRNA
A;Residues: 1-34,'R',36-154,168-264,'Y',266-321,'C',323-451,'F',453-581,'R',583-593,'H',5
A;Cross-references: GB:M23560; NID:g158795; PIDN:AAA29013.1; PID:g158796
A;Note: the authors translated the codon TGT for residue 322 as Val
C;Genetics:
A;Gene: vasa
A;Cross-references: FlyBase:FBgn0003970
A;Introns: 8/3; 138/1; 234/3; 482/3; 554/1; 644/2
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;289-296/Region: nucleotide-binding motif A (P-loop)
F;395-400/Region: nucleotide-binding motif B
F;399-402/Region: DEAD motif

Alignment Scores:
Pred. No.:              2.65e-14      Length:        661
Score:                  23.00         Matches:       23
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.18%         Indels:        0
DB:                     2             Gaps:          0

US-09-714-865-15 (1-2172) x A58768 (1-661)

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeu 301

Qy       1033 CCAATTTTG 1041
              |||||||||
Db        302 ProIleLeu 304

RESULT 6
A32378
probable ATP-dependent RNA helicase - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Sep-1989 #sequence_revision 08-Sep-1989 #text_change 19-Jan-2001
C;Accession: A32378
R;Leroy, P.; Alzari, P.; Sassoon, D.; Wolgemuth, D.; Fellous, M.
Cell 57, 549-559, 1989
A;Title: The protein encoded by a murine male germ cell-specific transcript is a putative
A;Reference number: A32378; MUID:89249320; PMID:2720782
A;Accession: A32378
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-660 <LER>
A;Cross-references: GB:J04847; NID:g200388; PIDN:AAA39942.1; PID:g200389
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; DNA binding; nucleotide binding; P-loop
F;223-230/Region: nucleotide-binding motif A (P-loop)
F;342-347/Region: nucleotide-binding motif B
F;346-349/Region: DEAD motif

Alignment Scores:
Pred. No.:              2.78e-13      Length:        660
Score:                  22.00         Matches:       22
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.04%         Indels:        0
DB:                     2             Gaps:          0

US-09-714-865-15 (1-2172) x A32378 (1-660)

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 236

Qy       1036 ATTTTG 1041
              ||||||
```

```
Db        237 IleLeu 238

RESULT 7
I84741
RNA helicase - mouse
N;Alternate names: RNA helicase ERH
C;Species: Mus musculus (house mouse)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 19-Jan-2001
C;Accession: I84741; S56112
R;Gee, S.L.; Conboy, J.G.
Gene 140, 171-177, 1994
A;Title: Mouse erythroid cells express multiple putative RNA helicase genes exhibiting h
A;Reference number: I49731; MUID:94192995; PMID:8144024
A;Accession: I84741
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-662 <RES>
A;Cross-references: GB:L25126; NID:g407995; PIDN:AAA53630.1; PID:g407996
A;Genetics: RES1
R;Sowden, J.; Putt, W.; Morrison, K.; Beddington, R.; Edwards, Y.
Biochem. J. 308, 839-846, 1995
A;Title: The embryonic RNA helicase gene (ERH): a new member of the DEAD box family of R
A;Reference number: S56112; MUID:97104282; PMID:8948440
A;Accession: S56112
A;Molecule type: mRNA
A;Residues: 1-229,'R',231-662 <SOW>
A;Cross-references: GB:Z38117; NID:g1835121; PID:g1835122
A;Experimental source: strain c57bl/6; notochord
A;Genetics: SOW1
A;Note: the sequence is revised in GenBank entry MMDBRNAHL, release 117, (PID:1835122)
A;Note: the revised sequence is now identical to PIR accession I84741
C;Genetics: <RES1>
A;Gene: Eif4a-rs1; MGI:Ddx19
A;Cross-references: MGI:99526
C;Genetics: <SOW1>
A;Gene: MGI:Ddx3
A;Cross-references: MGI:103064
A;Map position: 1
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;224-231/Region: nucleotide-binding motif A (P-loop)
F;343-348/Region: nucleotide-binding motif B
F;347-350/Region: DEAD motif

Alignment Scores:
Pred. No.:              2.78e-13      Length:        662
Score:                  22.00         Matches:       22
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.04%         Indels:        0
DB:                     1             Gaps:          0

US-09-714-865-15 (1-2172) x I84741 (1-662)

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 237

Qy       1036 ATTTTG 1041
              ||||||
Db        238 IleLeu 239

RESULT 8
S13654
ATP-dependent RNA helicase - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: S13654; S29676
R;Gururajan, R.; Perry-O'Keefe, H.; Melton, D.A.; Weeks, D.L.
Nature 349, 717-719, 1991
A;Title: The Xenopus localized messenger RNA An3 may encode an ATP-dependent RNA helicas
A;Reference number: S13654; MUID:91141586; PMID:1996140
A;Accession: S13654
A;Molecule type: mRNA
A;Residues: 1-697 <GUR>
A;Cross-references: EMBL:X57328; NID:g65059; PIDN:CAA40605.1; PID:g65060
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;260-272/Region: ATP binding #status predicted
F;265-272/Region: nucleotide-binding motif A (P-loop)
F;384-389/Region: nucleotide-binding motif B
F;386-398/Region: ATP binding #status predicted
F;388-391/Region: DEAD motif

Alignment Scores:
Pred. No.:              2.76e-13      Length:        697
Score:                  22.00         Matches:       22
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.04%         Indels:        0
DB:                     1             Gaps:          0

US-09-714-865-15 (1-2172) x S13654 (1-697)

Qy        976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 278

Qy       1036 ATTTTG 1041
              ||||||
Db        279 IleLeu 280

RESULT 9
T45677
ATP-dependent RNA helicase-like protein - Arabidopsis thaliana
N;Alternate names: protein F14P22.160
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
C;Accession: T45677
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23011
A;Accession: T45677
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-646 <DAN>
A;Cross-references: EMBL:AL137082
A;Experimental source: cultivar Columbia; BAC clone F14P22
C;Genetics:
A;Map position: 3
A;Introns: 239/3; 267/3; 348/3; 404/3; 442/3
A;Note: F14P22.160
C;Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:
Pred. No.:              3.22e-10      Length:        646
Score:                  19.00         Matches:       19
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.62%         Indels:        0
DB:                     2             Gaps:          0

US-09-714-865-15 (1-2172) x T45677 (1-646)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPhe 200

RESULT 10
C87818
protein glh-1 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: C87818
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
```

```
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: C87818
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-604 <STO>
A;Cross-references: GB:chr_I; PIDN:AAB52901.1; PID:g1947009; GSPDB:GN00019; CESP:T21G5.3
A;Note: T21G5.3
C;Genetics:
A;Gene: glh-1
A;Map position: 1
C;Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:
Pred. No.:               3.41e-09       Length:        604
Score:                   18.00          Matches:       18
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             2.49%          Indels:        0
DB:                      2              Gaps:          0

US-09-714-865-15 (1-2172) x C87818 (1-604)

Qy         985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         223 MetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIle 240

RESULT 11
T15132
ATP-dependent RNA helicase GLH-1 - Caenorhabditis elegans (fragment)
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999.#text_change 17-Nov-2000
C;Accession: T15132
R;Minx, P.
submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid T21G5.
A;Reference number: Z18299
A;Accession: T15132
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-604 <MIN>
A;Cross-references: EMBL:AF000197; NID:g1947004; PID:g1947009; PIDN:AAB52901.1; GSPDB:GN
A;Experimental source: strain Bristol N2; clone T21G5
C;Genetics:
A;Gene: CESP:glh-1
A;Map position: 1

Alignment Scores:
Pred. No.:               3.41e-09       Length:        604
Score:                   18.00          Matches:       18
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             2.49%          Indels:        0
DB:                      2              Gaps:          0

US-09-714-865-15 (1-2172) x T15132 (1-604)

Qy         985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         223 MetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIle 240

RESULT 12
H84854
probable ATP-dependent RNA helicase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: H84854
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
```

```
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84854
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-633 <STO>
A;Cross-references: GB:AE002093; NID:g4559339; PIDN:AAD23001.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g42520
A;Map position: 2
C;Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:
Pred. No.:               3.39e-09       Length:        633
Score:                   18.00          Matches:       18
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             2.49%          Indels:        0
DB:                      2              Gaps:          0

US-09-714-865-15 (1-2172) x H84854 (1-633)

Qy         973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         196 GlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPhe 213

RESULT 13
T45671
ATP-dependent RNA helicase-like protein - Arabidopsis thaliana
N;Alternate names: protein F14P22.100
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
C;Accession: T45671
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23011
A;Accession: T45671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-603 <DAN>
A;Cross-references: EMBL:AL137082
A;Experimental source: cultivar Columbia; BAC clone F14P22
C;Genetics:
A;Map position: 3
A;Introns: 244/3; 272/3; 353/3; 407/3; 436/3
A;Note: F14P22.100
C;Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:
Pred. No.:               3.58e-08       Length:        603
Score:                   17.00          Matches:       17
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             2.35%          Indels:        0
DB:                      2              Gaps:          0

US-09-714-865-15 (1-2172) x T45671 (1-603)

Qy         976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         189 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPhe 205

RESULT 14
S62003
probable ATP-dependent RNA helicase DBP1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: probable RNA helicase CA1; protein LPH8c; protein YPL119c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 19-Jan-2001
C;Accession: S62003; S16790; A34848
R;Schlenstedt, G.; Silver, P.A.
submitted to the EMBL Data Library, December 1995
A;Reference number: S61996
A;Accession: S62003
```

```
A;Molecule type: DNA
A;Residues: 1-617 <SCH>
A;Cross-references: EMBL:U43503; NID:g1163087; PIDN:AAB68243.1; PID:g1163095; MIPS:YPL11
R;Jamieson, D.J.; Beggs, J.D.
Mol. Microbiol. 5, 805-812, 1991
A;Title: A suppressor of yeast spp81/ded1 mutations encodes a very similar putative ATP-
A;Reference number: S16790; MUID:91312117; PMID:1857205
A;Accession: S16790
A;Molecule type: DNA
A;Residues: 1-42,'RS',45-47,'K',49-87,'R',89-114,'QK',116-617 <JAM>
A;Cross-references: EMBL:X55993; NID:g3640; PIDN:CAA39465.1; PID:g3641
A;Experimental source: strain DBY939
R;Chang, T.H.; Arenas, J.; Abelson, J.
Proc. Natl. Acad. Sci. U.S.A. 87, 1571-1575, 1990
A;Title: Identification of five putative yeast RNA helicase genes.
A;Reference number: A34848; MUID:90160368; PMID:2406722
A;Accession: A34848
A;Molecule type: DNA
A;Residues: 316-495,'I',497-500 <CHA>
C;Genetics:
A;Gene: SGD:DBP1
A;Cross-references: SGD:S0006040; MIPS:YPL119c
A;Map position: 16L
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;198-205/Region: nucleotide-binding motif A (P-loop)
F;314-319/Region: nucleotide-binding motif B
F;318-321/Region: DEAD motif

Alignment Scores:
Pred. No.:              3.57e-08      Length:        617
Score:                  17.00         Matches:       17
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.35%         Indels:        0
DB:                     2             Gaps:          0

US-09-714-865-15 (1-2172) x S62003 (1-617)

Qy       1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 IleLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 328

RESULT 15
T48796
probable ATP-dependent RNA helicase DED1 [imported] - Neurospora crassa
N;Alternate names: protein 15E6.40
C;Species: Neurospora crassa
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 17-Nov-2000
C;Accession: T48796
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24541
A;Accession: T48796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-688 <SCH>
A;Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15E6.40
A;Experimental source: cosmid contig 15E6; strain 74
C;Genetics:
A;Gene: NCSP:15E6.40
A;Map position: 2
A;Introns: 63/2
C;Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:
Pred. No.:              3.51e-08      Length:        688
Score:                  17.00         Matches:       17
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.35%         Indels:        0
DB:                     2             Gaps:          0

US-09-714-865-15 (1-2172) x T48796 (1-688)

Qy       1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 IleLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 370

Search completed: June 10, 2003, 17:10:06
Job time : 73 secs
```

OM nucleic - protein search, using frame_plus_n2p model

```
Run on:         June 10, 2003, 17:08:46 ; Search time 42.5 Seconds
                                          (without alignments)
                                          10552.370 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  724
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  OLIGO
                Xgapop 60.0 , Xgapext 60.0
                Ygapop 60.0 , Ygapext 60.0
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       392085 seqs, 103240269 residues

Word size:      1

Total number of hits satisfying chosen parameters:      738648

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111835_26709/app_query.fasta_1.2311
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=oligo.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09714865_@CGN_1_1_59_@runat_05062003_111835_26709
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :      Published_Applications_AA:*
                1:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                10:  /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                12:  /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
                13:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                14:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	84	11.6	106	9	US-09-764-891-3857	Sequence 3857, Ap
	2	15	2.1	238	9	US-10-195-117-3	Sequence 3, Appli
	3	15	2.1	238	9	US-10-195-117-4	Sequence 4, Appli
	4	15	2.1	343	10	US-09-925-301-1439	Sequence 1439, Ap
	5	15	2.1	500	9	US-10-228-897-13	Sequence 13, Appl
	6	15	2.1	575	9	US-10-108-605-55	Sequence 55, Appl
	7	11	1.5	45	10	US-09-864-761-46256	Sequence 46256, A
	8	11	1.5	613	10	US-09-815-242-10990	Sequence 10990, A
	9	11	1.5	646	10	US-09-815-242-10311	Sequence 10311, A
	10	11	1.5	646	10	US-09-815-242-13769	Sequence 13769, A
	11	11	1.5	648	10	US-09-923-831-43	Sequence 43, Appl
	12	10	1.4	421	10	US-09-815-242-11703	Sequence 11703, A
	13	10	1.4	424	9	US-09-738-626-4359	Sequence 4359, Ap
	14	10	1.4	495	9	US-09-971-536-52	Sequence 52, Appl
	15	10	1.4	524	10	US-09-815-242-13491	Sequence 13491, A
	16	10	1.4	536	10	US-09-815-242-10774	Sequence 10774, A
	17	10	1.4	567	10	US-09-815-242-11885	Sequence 11885, A
	18	10	1.4	624	10	US-09-963-790A-2	Sequence 2, Appli
	19	10	1.4	732	9	US-09-738-626-4778	Sequence 4778, Ap
c	20	9	1.3	338	9	US-09-953-280-28	Sequence 28, Appl
c	21	9	1.3	338	9	US-09-953-280-29	Sequence 29, Appl
c	22	9	1.3	338	9	US-10-040-949A-44	Sequence 44, Appl
c	23	9	1.3	338	9	US-10-040-949A-45	Sequence 45, Appl
c	24	9	1.3	543	9	US-09-736-457-337	Sequence 337, App
c	25	9	1.3	543	9	US-09-902-941-337	Sequence 337, App
c	26	9	1.3	543	9	US-09-849-626-337	Sequence 337, App
c	27	9	1.3	543	9	US-10-017-754-337	Sequence 337, App
	28	9	1.2	709	9	US-09-736-457-335	Sequence 335, App
	29	9	1.2	709	9	US-09-902-941-335	Sequence 335, App
	30	9	1.2	709	9	US-09-849-626-335	Sequence 335, App
	31	9	1.2	709	9	US-10-017-754-335	Sequence 335, App
	32	9	1.2	1261	12	US-10-147-268-2	Sequence 2, Appli
c	33	8	1.1	16	10	US-09-835-087-14	Sequence 14, Appl
c	34	8	1.1	39	10	US-09-864-761-42449	Sequence 42449, A
	35	8	1.1	41	10	US-09-732-091-12	Sequence 12, Appl
c	36	8	1.1	58	9	US-10-091-504-1215	Sequence 1215, Ap
c	37	8	1.1	58	10	US-09-764-869-1215	Sequence 1215, Ap
c	38	8	1.1	64	9	US-09-764-891-4988	Sequence 4988, Ap
c	39	8	1.1	66	9	US-09-984-271-147	Sequence 147, App
c	40	8	1.1	72	10	US-09-729-674-58	Sequence 58, Appl
c	41	8	1.1	95	9	US-09-984-271-234	Sequence 234, App
c	42	8	1.1	100	10	US-09-840-459-21	Sequence 21, Appl
c	43	8	1.1	100	10	US-09-840-459-22	Sequence 22, Appl
c	44	8	1.1	101	9	US-10-194-975-73	Sequence 73, Appl
c	45	8	1.1	101	9	US-10-194-975-74	Sequence 74, Appl
c	46	8	1.1	110	10	US-09-864-761-48276	Sequence 48276, A
c	47	8	1.1	111	9	US-09-995-529-14	Sequence 14, Appl
c	48	8	1.1	112	9	US-10-032-482-8	Sequence 8, Appli
c	49	8	1.1	112	10	US-09-772-120-6	Sequence 6, Appli
c	50	8	1.1	112	10	US-09-835-087-1	Sequence 1, Appli
c	51	8	1.1	112	10	US-09-835-087-3	Sequence 3, Appli
c	52	8	1.1	112	10	US-09-835-087-4	Sequence 4, Appli
c	53	8	1.1	112	10	US-09-835-087-5	Sequence 5, Appli
c	54	8	1.1	112	10	US-09-835-087-6	Sequence 6, Appli
c	55	8	1.1	112	10	US-09-835-087-7	Sequence 7, Appli
c	56	8	1.1	112	10	US-09-809-739-11	Sequence 11, Appl
c	57	8	1.1	112	10	US-09-809-739-14	Sequence 14, Appl
c	58	8	1.1	112	10	US-09-809-739-15	Sequence 15, Appl
c	59	8	1.1	112	10	US-09-809-739-16	Sequence 16, Appl
c	60	8	1.1	112	10	US-09-809-739-17	Sequence 17, Appl
c	61	8	1.1	112	10	US-09-809-739-18	Sequence 18, Appl
c	62	8	1.1	112	10	US-09-840-459-9	Sequence 9, Appli
c	63	8	1.1	112	10	US-09-840-459-12	Sequence 12, Appl
c	64	8	1.1	112	10	US-09-840-459-13	Sequence 13, Appl
c	65	8	1.1	112	10	US-09-840-459-14	Sequence 14, Appl
c	66	8	1.1	112	10	US-09-840-459-15	Sequence 15, Appl
c	67	8	1.1	112	10	US-09-840-459-54	Sequence 54, Appl
c	68	8	1.1	112	10	US-09-840-459-107	Sequence 107, App
c	69	8	1.1	113	9	US-09-995-529-18	Sequence 18, Appl
c	70	8	1.1	113	10	US-09-840-459-68	Sequence 68, Appl
c	71	8	1.1	114	10	US-09-840-459-106	Sequence 106, App
c	72	8	1.1	133	10	US-09-925-301-1176	Sequence 1176, Ap
c	73	8	1.1	142	10	US-09-840-459-102	Sequence 102, App
	74	8	1.1	160	9	US-09-991-496-135	Sequence 135, App
	75	8	1.1	211	10	US-09-925-297-823	Sequence 823, App
	76	8	1.1	220	9	US-09-925-299-843	Sequence 843, App
	77	8	1.1	220	10	US-09-925-299-843	Sequence 843, App

c	78	8	1.1	233	10	US-09-938-803-15	Sequence 15, Appl
c	79	8	1.1	236	10	US-09-925-300-1410	Sequence 1410, Ap
	80	8	1.1	239	9	US-10-195-117-6	Sequence 6, Appl
c	81	8	1.1	271	10	US-09-911-346-2	Sequence 2, Appli
c	82	8	1.1	320	9	US-10-278-173-68	Sequence 68, Appl
	83	8	1.1	328	10	US-09-815-242-13488	Sequence 13488, A
	84	8	1.1	328	10	US-09-815-242-13556	Sequence 13556, A
c	85	8	1.1	353	9	US-09-953-280-42	Sequence 42, Appl
c	86	8	1.1	353	9	US-10-040-949A-58	Sequence 58, Appl
	87	8	1.1	387	10	US-09-925-300-1477	Sequence 1477, Ap
c	88	8	1.1	440	9	US-09-738-626-6376	Sequence 6376, Ap
	89	8	1.1	459	10	US-09-925-301-1027	Sequence 1027, Ap
	90	8	1.1	471	9	US-10-082-830-265	Sequence 265, App
	91	8	1.1	487	10	US-09-815-242-12397	Sequence 12397, A
	92	8	1.1	487	10	US-09-815-242-12991	Sequence 12991, A
	93	8	1.1	492	10	US-09-815-242-11308	Sequence 11308, A
	94	8	1.1	492	10	US-09-732-091-2	Sequence 2, Appli
	95	8	1.1	504	10	US-09-732-091-42	Sequence 42, Appl
	96	8	1.1	510	10	US-09-815-242-5256	Sequence 5256, Ap
	97	8	1.1	526	9	US-10-084-205-74	Sequence 74, Appl
	98	8	1.1	526	10	US-09-925-637-74	Sequence 74, Appl
c	99	8	1.1	535	9	US-09-968-851-38	Sequence 38, Appl
	100	8	1.1	548	9	US-10-023-282-264	Sequence 264, App
	101	8	1.1	633	9	US-09-834-998A-3	Sequence 3, Appli
	102	8	1.1	725	9	US-10-108-605-321	Sequence 321, App
	103	8	1.1	752	10	US-09-825-809-2	Sequence 2, Appli
	104	8	1.1	752	10	US-09-825-809-4	Sequence 4, Appli
	105	8	1.1	764	10	US-09-902-627-1	Sequence 1, Appli
c	106	8	1.1	805	9	US-10-108-605-113	Sequence 113, App
c	107	8	1.1	1345	9	US-10-108-605-249	Sequence 249, App
c	108	8	1.1	1661	9	US-09-842-758-42	Sequence 42, Appl
c	109	8	1.1	1744	9	US-10-108-605-25	Sequence 25, Appl
	110	7	1.0	10	9	US-09-572-404B-1213	Sequence 1213, Ap
	111	7	1.0	12	9	US-10-057-789-305	Sequence 305, App
	112	7	1.0	12	9	US-10-212-628-305	Sequence 305, App
c	113	7	1.0	12	9	US-09-954-385-134	Sequence 134, App
c	114	7	1.0	14	9	US-09-984-245-147	Sequence 147, App
c	115	7	1.0	14	9	US-09-966-262-147	Sequence 147, App
c	116	7	1.0	14	9	US-09-983-966-147	Sequence 147, App
c	117	7	1.0	14	9	US-10-143-090-147	Sequence 147, App
	118	7	1.0	18	9	US-09-759-130B-294	Sequence 294, App
	119	7	1.0	18	9	US-10-189-123-24	Sequence 24, Appl
	120	7	1.0	18	10	US-09-864-761-40388	Sequence 40388, A
	121	7	1.0	19	10	US-09-879-957-170	Sequence 170, App
c	122	7	1.0	22	9	US-10-097-065-369	Sequence 369, App
	123	7	1.0	22	10	US-09-071-838-302	Sequence 302, App
c	124	7	1.0	28	10	US-09-864-761-33674	Sequence 33674, A
c	125	7	1.0	31	10	US-09-864-761-48994	Sequence 48994, A
c	126	7	1.0	33	9	US-10-042-141-155	Sequence 155, App
c	127	7	1.0	33	10	US-09-726-643-155	Sequence 155, App
c	128	7	1.0	34	10	US-09-864-761-34605	Sequence 34605, A
	129	7	1.0	34	10	US-09-864-761-39859	Sequence 39859, A
c	130	7	1.0	37	10	US-09-864-761-42324	Sequence 42324, A
c	131	7	1.0	41	10	US-09-864-761-37109	Sequence 37109, A
c	132	7	1.0	43	9	US-09-981-876-205	Sequence 205, App
c	133	7	1.0	43	9	US-09-148-545-205	Sequence 205, App
c	134	7	1.0	45	9	US-10-072-349-102	Sequence 102, App
c	135	7	1.0	45	10	US-09-764-855-102	Sequence 102, App
c	136	7	1.0	48	9	US-09-984-271-165	Sequence 165, App
	137	7	1.0	53	9	US-10-150-111-81	Sequence 81, Appl
c	138	7	1.0	56	9	US-10-114-893-44	Sequence 44, Appl
c	139	7	1.0	56	9	US-10-114-893-129	Sequence 129, App
c	140	7	1.0	60	10	US-09-864-761-40431	Sequence 40431, A
c	141	7	1.0	61	9	US-09-983-802-621	Sequence 621, App
c	142	7	1.0	64	10	US-09-922-261-267	Sequence 267, App
	143	7	1.0	65	10	US-09-867-550-54	Sequence 54, Appl
c	144	7	1.0	72	9	US-09-746-783-118	Sequence 118, App
c	145	7	1.0	75	10	US-09-922-261-265	Sequence 265, App
	146	7	1.0	76	9	US-10-012-896-575	Sequence 575, App
	147	7	1.0	76	9	US-10-012-896-888	Sequence 888, App
	148	7	1.0	76	9	US-09-895-793-575	Sequence 575, App
	149	7	1.0	76	9	US-09-895-793-888	Sequence 888, App
	150	7	1.0	76	9	US-09-895-814-575	Sequence 575, App

	151	7	1.0	76	9	US-09-895-814-888	Sequence 888, App
	152	7	1.0	76	10	US-09-759-143-575	Sequence 575, App
	153	7	1.0	76	10	US-09-759-143-888	Sequence 888, App
	154	7	1.0	76	10	US-09-780-669-575	Sequence 575, App
	155	7	1.0	76	10	US-09-780-669-888	Sequence 888, App
	156	7	1.0	76	10	US-09-822-827-575	Sequence 575, App
	157	7	1.0	76	10	US-09-822-827-888	Sequence 888, App
c	158	7	1.0	77	10	US-09-864-761-37461	Sequence 37461, A
c	159	7	1.0	80	10	US-09-864-761-46876	Sequence 46876, A
c	160	7	1.0	86	10	US-09-864-761-34118	Sequence 34118, A
	161	7	1.0	88	10	US-09-939-980-513	Sequence 513, App
c	162	7	1.0	89	9	US-09-764-891-3157	Sequence 3157, Ap
c	163	7	1.0	95	10	US-09-864-761-48467	Sequence 48467, A
c	164	7	1.0	97	10	US-09-734-017A-40	Sequence 40, Appl
c	165	7	1.0	98	10	US-09-864-761-34193	Sequence 34193, A
	166	7	1.0	99	10	US-09-815-242-5533	Sequence 5533, Ap
c	167	7	1.0	101	9	US-10-097-065-366	Sequence 366, App
c	168	7	1.0	105	9	US-09-986-480-231	Sequence 231, App
	169	7	1.0	105	10	US-09-815-242-12217	Sequence 12217, A
	170	7	1.0	105	10	US-09-815-242-12760	Sequence 12760, A
c	171	7	1.0	111	10	US-09-893-737-122	Sequence 122, App
c	172	7	1.0	111	10	US-09-746-491-16	Sequence 16, Appl
	173	7	1.0	112	9	US-09-898-751A-12	Sequence 12, Appl
	174	7	1.0	112	9	US-10-146-496-6	Sequence 6, Appli
	175	7	1.0	112	10	US-09-931-381A-11	Sequence 11, Appl
c	176	7	1.0	114	9	US-10-114-893-137	Sequence 137, App
c	177	7	1.0	117	9	US-10-193-616-6	Sequence 6, Appli
c	178	7	1.0	119	9	US-10-050-704-227	Sequence 227, App
c	179	7	1.0	122	10	US-09-893-737-286	Sequence 286, App
	180	7	1.0	123	9	US-09-796-692-1360	Sequence 1360, Ap
	181	7	1.0	123	9	US-09-796-692-2329	Sequence 2329, Ap
	182	7	1.0	123	9	US-10-040-862-1360	Sequence 1360, Ap
	183	7	1.0	123	9	US-10-040-862-2329	Sequence 2329, Ap
c	184	7	1.0	130	9	US-09-948-783-142	Sequence 142, App
c	185	7	1.0	131	9	US-09-892-877-140	Sequence 140, App
c	186	7	1.0	135	10	US-09-881-823-10	Sequence 10, Appl
c	187	7	1.0	136	9	US-10-050-704-107	Sequence 107, App
	188	7	1.0	137	9	US-09-764-891-4420	Sequence 4420, Ap
c	189	7	1.0	138	9	US-10-042-141-151	Sequence 151, App
c	190	7	1.0	138	10	US-09-726-643-151	Sequence 151, App
c	191	7	1.0	141	10	US-09-893-737-254	Sequence 254, App
c	192	7	1.0	142	9	US-09-966-546-22	Sequence 22, Appl
c	193	7	1.0	142	9	US-09-966-545-22	Sequence 22, Appl
c	194	7	1.0	142	9	US-09-965-212-22	Sequence 22, Appl
c	195	7	1.0	142	10	US-09-800-729-139	Sequence 139, App
c	196	7	1.0	142	10	US-09-800-729-198	Sequence 198, App
c	197	7	1.0	143	9	US-09-782-974C-62	Sequence 62, Appl
c	198	7	1.0	150	10	US-09-893-737-230	Sequence 230, App
c	199	7	1.0	152	9	US-09-764-891-4604	Sequence 4604, Ap
	200	7	1.0	155	9	US-09-738-626-4054	Sequence 4054, Ap
c	201	7	1.0	157	9	US-10-080-960-27	Sequence 27, Appl
c	202	7	1.0	161	9	US-10-099-766-2	Sequence 2, Appli
c	203	7	1.0	161	10	US-09-682-706-5	Sequence 5, Appli
	204	7	1.0	168	9	US-10-267-718-19	Sequence 19, Appl
	205	7	1.0	168	10	US-09-354-453-19	Sequence 19, Appl
c	206	7	1.0	175	9	US-10-050-704-228	Sequence 228, App
c	207	7	1.0	176	9	US-10-193-616-8	Sequence 8, Appli
c	208	7	1.0	176	10	US-09-855-266A-1	Sequence 1, Appli
	209	7	1.0	177	10	US-09-755-325-4	Sequence 4, Appli
c	210	7	1.0	180	9	US-09-989-442-107	Sequence 107, App
c	211	7	1.0	180	9	US-10-073-865-86	Sequence 86, Appl
c	212	7	1.0	180	9	US-10-103-313-379	Sequence 379, App
c	213	7	1.0	180	9	US-10-073-885-99	Sequence 99, Appl
c	214	7	1.0	180	10	US-09-764-853-567	Sequence 567, App
c	215	7	1.0	183	9	US-09-989-442-111	Sequence 111, App
c	216	7	1.0	183	9	US-10-073-865-95	Sequence 95, Appl
c	217	7	1.0	183	9	US-10-103-313-389	Sequence 389, App
c	218	7	1.0	183	9	US-10-073-885-102	Sequence 102, App
c	219	7	1.0	183	10	US-09-764-853-592	Sequence 592, App
	220	7	1.0	184	10	US-09-925-301-847	Sequence 847, App
c	221	7	1.0	185	10	US-09-529-063-34	Sequence 34, Appl
c	222	7	1.0	188	9	US-09-866-050A-686	Sequence 686, App
	223	7	1.0	198	9	US-09-892-877-212	Sequence 212, App

	224	7	1.0	198	9	US-09-948-783-213	Sequence 213, App
c	225	7	1.0	203	9	US-10-050-704-244	Sequence 244, App
c	226	7	1.0	210	9	US-09-973-025-14	Sequence 14, Appl
c	227	7	1.0	210	9	US-09-899-303-14	Sequence 14, Appl
c	228	7	1.0	210	9	US-09-995-808-14	Sequence 14, Appl
c	229	7	1.0	215	9	US-10-255-532-4	Sequence 4, Appli
c	230	7	1.0	223	9	US-09-875-321-9	Sequence 9, Appli
c	231	7	1.0	223	9	US-10-162-012-9	Sequence 9, Appli
c	232	7	1.0	228	10	US-09-780-717-29	Sequence 29, Appl
c	233	7	1.0	229	10	US-09-804-357-12	Sequence 12, Appl
c	234	7	1.0	229	10	US-09-804-006-12	Sequence 12, Appl
c	235	7	1.0	242	10	US-09-864-761-36180	Sequence 36180, A
	236	7	1.0	242	10	US-09-867-550-1878	Sequence 1878, Ap
c	237	7	1.0	245	9	US-10-128-870-8	Sequence 8, Appli
c	238	7	1.0	245	9	US-10-131-685-8	Sequence 8, Appli
	239	7	1.0	245	10	US-09-942-858-9	Sequence 9, Appli
c	240	7	1.0	247	10	US-09-893-348-14	Sequence 14, Appl
c	241	7	1.0	248	9	US-10-233-873A-3	Sequence 3, Appli
	242	7	1.0	257	8	US-08-450-842-10	Sequence 10, Appl
c	243	7	1.0	259	12	US-10-078-929-197	Sequence 197, App
c	244	7	1.0	266	10	US-09-911-826A-6	Sequence 6, Appli
	245	7	1.0	268	10	US-09-864-866-43	Sequence 43, Appl
c	246	7	1.0	272	9	US-10-102-806-663	Sequence 663, App
	247	7	1.0	275	9	US-10-112-645-4	Sequence 4, Appli
c	248	7	1.0	277	9	US-10-101-464A-631	Sequence 631, App
c	249	7	1.0	284	9	US-10-073-885-71	Sequence 71, Appl
	250	7	1.0	284	10	US-09-771-161A-106	Sequence 106, App
	251	7	1.0	290	9	US-10-117-323-4	Sequence 4, Appli
c	252	7	1.0	292	9	US-10-278-173-18	Sequence 18, Appl
c	253	7	1.0	292	10	US-09-864-761-37944	Sequence 37944, A
c	254	7	1.0	293	10	US-09-815-242-5614	Sequence 5614, Ap
c	255	7	1.0	298	10	US-09-789-919-68	Sequence 68, Appl
c	256	7	1.0	299	9	US-10-125-001-22	Sequence 22, Appl
c	257	7	1.0	300	9	US-10-128-870-4	Sequence 4, Appli
c	258	7	1.0	300	9	US-10-128-870-6	Sequence 6, Appli
c	259	7	1.0	300	9	US-10-131-685-4	Sequence 4, Appli
c	260	7	1.0	300	9	US-10-131-685-6	Sequence 6, Appli
	261	7	1.0	300	10	US-09-861-451A-42	Sequence 42, Appl
c	262	7	1.0	301	9	US-10-004-633-24	Sequence 24, Appl
	263	7	1.0	301	9	US-10-183-116-71	Sequence 71, Appl
c	264	7	1.0	303	9	US-09-828-523A-20	Sequence 20, Appl
c	265	7	1.0	303	9	US-09-533-029-6	Sequence 6, Appli
c	266	7	1.0	303	10	US-09-815-242-12527	Sequence 12527, A
c	267	7	1.0	308	9	US-10-097-065-150	Sequence 150, App
	268	7	1.0	309	9	US-09-991-496-134	Sequence 134, App
c	269	7	1.0	311	9	US-09-828-523A-86	Sequence 86, Appl
	270	7	1.0	313	9	US-09-771-503-9	Sequence 9, Appli
c	271	7	1.0	326	10	US-09-804-357-8	Sequence 8, Appli
c	272	7	1.0	326	10	US-09-804-006-8	Sequence 8, Appli
c	273	7	1.0	332	9	US-10-236-250-2	Sequence 2, Appli
c	274	7	1.0	332	9	US-09-804-291-63	Sequence 63, Appl
c	275	7	1.0	332	10	US-09-767-041-21	Sequence 21, Appl
c	276	7	1.0	332	10	US-09-886-055-63	Sequence 63, Appl
	277	7	1.0	332	10	US-09-764-864-1573	Sequence 1573, Ap
c	278	7	1.0	332	10	US-09-771-730-10	Sequence 10, Appl
c	279	7	1.0	332	10	US-09-771-730-28	Sequence 28, Appl
	280	7	1.0	340	9	US-09-738-626-5037	Sequence 5037, Ap
c	281	7	1.0	344	10	US-09-866-582-6	Sequence 6, Appli
c	282	7	1.0	348	10	US-09-815-242-5130	Sequence 5130, Ap
	283	7	1.0	353	9	US-09-759-130B-283	Sequence 283, App
	284	7	1.0	353	9	US-10-189-123-13	Sequence 13, Appl
c	285	7	1.0	359	9	US-10-093-246-1	Sequence 1, Appli
c	286	7	1.0	359	12	US-10-093-045-1	Sequence 1, Appli
c	287	7	1.0	360	10	US-09-925-301-1116	Sequence 1116, Ap
c	288	7	1.0	361	10	US-09-796-858-28	Sequence 28, Appl
c	289	7	1.0	362	9	US-10-093-246-2	Sequence 2, Appli
c	290	7	1.0	362	12	US-10-093-045-2	Sequence 2, Appli
	291	7	1.0	366	10	US-09-934-899-6	Sequence 6, Appli
	292	7	1.0	366	10	US-09-934-868-26	Sequence 26, Appl
c	293	7	1.0	370	9	US-10-097-065-381	Sequence 381, App
c	294	7	1.0	370	10	US-09-893-737-202	Sequence 202, App
	295	7	1.0	371	9	US-09-923-844B-2	Sequence 2, Appli
	296	7	1.0	371	9	US-09-759-130B-281	Sequence 281, App
	297	7	1.0	371	9	US-10-189-123-11	Sequence 11, Appl
c	298	7	1.0	373	9	US-10-318-142-6	Sequence 6, Appli
c	299	7	1.0	373	10	US-09-760-354A-2	Sequence 2, Appli
	300	7	1.0	373	10	US-09-820-893-55	Sequence 55, Appl
c	301	7	1.0	375	10	US-09-764-864-1316	Sequence 1316, Ap
c	302	7	1.0	378	9	US-10-073-885-77	Sequence 77, Appl
c	303	7	1.0	378	10	US-09-764-864-863	Sequence 863, App
	304	7	1.0	382	12	US-10-078-929-34	Sequence 34, Appl
	305	7	1.0	383	9	US-09-738-626-6821	Sequence 6821, Ap
c	306	7	1.0	386	9	US-10-270-333-18	Sequence 18, Appl
	307	7	1.0	387	9	US-10-114-893-133	Sequence 133, App
	308	7	1.0	387	9	US-10-016-249-2	Sequence 2, Appli
c	309	7	1.0	393	10	US-09-823-240-11	Sequence 11, Appl
c	310	7	1.0	396	9	US-10-193-616-14	Sequence 14, Appl
c	311	7	1.0	399	10	US-09-893-737-10	Sequence 10, Appl
	312	7	1.0	402	10	US-09-764-864-1496	Sequence 1496, Ap
	313	7	1.0	403	9	US-09-991-496-10	Sequence 10, Appl
	314	7	1.0	403	10	US-09-874-923-10	Sequence 10, Appl
c	315	7	1.0	404	9	US-09-712-363-257	Sequence 257, App
	316	7	1.0	404	10	US-09-764-864-1074	Sequence 1074, Ap
c	317	7	1.0	407	9	US-09-770-509-16	Sequence 16, Appl
	318	7	1.0	408	9	US-09-988-462-21	Sequence 21, Appl
	319	7	1.0	410	10	US-09-925-300-1074	Sequence 1074, Ap
c	320	7	1.0	411	1	US-08-954-771-10	Sequence 10, Appl
c	321	7	1.0	411	8	US-08-900-220C-12	Sequence 12, Appl
c	322	7	1.0	411	9	US-09-883-848A-12	Sequence 12, Appl
c	323	7	1.0	411	9	US-09-187-387-12	Sequence 12, Appl
	324	7	1.0	411	10	US-09-755-325-6	Sequence 6, Appli
	325	7	1.0	411	10	US-09-755-325-8	Sequence 8, Appli
c	326	7	1.0	411	10	US-09-151-999-12	Sequence 12, Appl
	327	7	1.0	415	10	US-09-755-325-2	Sequence 2, Appli
c	328	7	1.0	417	9	US-10-075-846-2	Sequence 2, Appli
c	329	7	1.0	431	9	US-10-075-846-4	Sequence 4, Appli
c	330	7	1.0	434	10	US-09-801-368-146	Sequence 146, App
	331	7	1.0	437	9	US-09-941-947A-2	Sequence 2, Appli
	332	7	1.0	437	10	US-09-815-242-5233	Sequence 5233, Ap
	333	7	1.0	437	10	US-09-934-901-16	Sequence 16, Appl
	334	7	1.0	437	10	US-09-934-868-6	Sequence 6, Appli
c	335	7	1.0	440	9	US-10-063-547-134	Sequence 134, App
c	336	7	1.0	440	9	US-10-174-590-386	Sequence 386, App
c	337	7	1.0	440	9	US-10-176-758-386	Sequence 386, App
c	338	7	1.0	440	9	US-10-063-616-134	Sequence 134, App
c	339	7	1.0	440	9	US-10-175-737-386	Sequence 386, App
c	340	7	1.0	440	9	US-10-063-502-134	Sequence 134, App
c	341	7	1.0	440	9	US-10-173-706-386	Sequence 386, App
c	342	7	1.0	440	9	US-10-175-738-386	Sequence 386, App
c	343	7	1.0	440	9	US-10-175-752-386	Sequence 386, App
c	344	7	1.0	440	9	US-10-176-482-386	Sequence 386, App
c	345	7	1.0	440	9	US-10-176-757-386	Sequence 386, App
c	346	7	1.0	440	9	US-10-176-913-386	Sequence 386, App
c	347	7	1.0	440	9	US-10-180-552-386	Sequence 386, App
c	348	7	1.0	440	9	US-10-180-557-386	Sequence 386, App
c	349	7	1.0	440	9	US-10-173-700-386	Sequence 386, App
c	350	7	1.0	440	9	US-10-174-572-386	Sequence 386, App
c	351	7	1.0	440	9	US-10-174-579-386	Sequence 386, App
c	352	7	1.0	440	9	US-10-174-582-386	Sequence 386, App
c	353	7	1.0	440	9	US-10-174-588-386	Sequence 386, App
c	354	7	1.0	440	9	US-10-175-739-386	Sequence 386, App
c	355	7	1.0	440	9	US-10-175-740-386	Sequence 386, App
c	356	7	1.0	440	9	US-10-175-743-386	Sequence 386, App
c	357	7	1.0	440	9	US-10-176-488-386	Sequence 386, App
c	358	7	1.0	440	9	US-10-176-492-386	Sequence 386, App
c	359	7	1.0	440	9	US-10-176-747-386	Sequence 386, App
c	360	7	1.0	440	9	US-10-176-750-386	Sequence 386, App
c	361	7	1.0	440	9	US-10-176-985-386	Sequence 386, App
c	362	7	1.0	440	9	US-10-176-987-386	Sequence 386, App
c	363	7	1.0	440	9	US-10-176-991-386	Sequence 386, App
c	364	7	1.0	440	9	US-10-176-992-386	Sequence 386, App
c	365	7	1.0	440	9	US-10-176-993-386	Sequence 386, App
c	366	7	1.0	440	9	US-10-184-658-386	Sequence 386, App
c	367	7	1.0	440	9	US-10-173-695-386	Sequence 386, App
c	368	7	1.0	440	9	US-10-173-697-386	Sequence 386, App
c	369	7	1.0	440	9	US-10-173-705-386	Sequence 386, App

c	370	7	1.0	440	9	US-10-174-576-386 Sequence 386, App
c	371	7	1.0	440	9	US-10-174-585-386 Sequence 386, App
c	372	7	1.0	440	9	US-10-174-586-386 Sequence 386, App
c	373	7	1.0	440	9	US-10-175-747-386 Sequence 386, App
c	374	7	1.0	440	9	US-10-176-481-386 Sequence 386, App
c	375	7	1.0	440	9	US-10-176-485-386 Sequence 386, App
c	376	7	1.0	440	9	US-10-176-487-386 Sequence 386, App
c	377	7	1.0	440	9	US-10-176-493-386 Sequence 386, App
c	378	7	1.0	440	9	US-10-176-756-386 Sequence 386, App
c	379	7	1.0	440	9	US-10-176-911-386 Sequence 386, App
c	380	7	1.0	440	9	US-10-176-919-386 Sequence 386, App
c	381	7	1.0	440	9	US-10-176-925-386 Sequence 386, App
c	382	7	1.0	440	9	US-10-176-978-386 Sequence 386, App
c	383	7	1.0	440	9	US-10-179-510-386 Sequence 386, App
c	384	7	1.0	440	9	US-10-180-543-386 Sequence 386, App
c	385	7	1.0	440	9	US-10-180-544-386 Sequence 386, App
c	386	7	1.0	440	9	US-10-180-546-386 Sequence 386, App
c	387	7	1.0	440	9	US-10-180-547-386 Sequence 386, App
c	388	7	1.0	440	9	US-10-180-549-386 Sequence 386, App
c	389	7	1.0	440	9	US-10-180-555-386 Sequence 386, App
c	390	7	1.0	440	9	US-10-180-559-386 Sequence 386, App
c	391	7	1.0	440	9	US-10-181-000-386 Sequence 386, App
c	392	7	1.0	440	9	US-10-183-010-386 Sequence 386, App
c	393	7	1.0	440	9	US-10-183-012-386 Sequence 386, App
c	394	7	1.0	440	9	US-10-184-614-386 Sequence 386, App
c	395	7	1.0	440	9	US-10-184-623-386 Sequence 386, App
c	396	7	1.0	440	9	US-10-184-635-386 Sequence 386, App
c	397	7	1.0	440	9	US-10-184-637-386 Sequence 386, App
c	398	7	1.0	440	9	US-10-184-646-386 Sequence 386, App
c	399	7	1.0	440	9	US-10-184-647-386 Sequence 386, App
c	400	7	1.0	440	9	US-10-184-652-386 Sequence 386, App
c	401	7	1.0	440	9	US-10-187-594-386 Sequence 386, App
c	402	7	1.0	440	9	US-10-187-596-386 Sequence 386, App
c	403	7	1.0	440	9	US-10-187-745-386 Sequence 386, App
c	404	7	1.0	440	9	US-10-187-885-386 Sequence 386, App
c	405	7	1.0	440	9	US-10-187-886-386 Sequence 386, App
c	406	7	1.0	440	9	US-10-199-464-386 Sequence 386, App
c	407	7	1.0	440	9	US-10-176-751-386 Sequence 386, App
c	408	7	1.0	440	9	US-10-176-760-386 Sequence 386, App
c	409	7	1.0	440	9	US-10-176-990-386 Sequence 386, App
c	410	7	1.0	440	9	US-10-180-541-386 Sequence 386, App
c	411	7	1.0	440	9	US-10-180-542-386 Sequence 386, App
c	412	7	1.0	440	9	US-10-180-548-386 Sequence 386, App
c	413	7	1.0	440	9	US-10-180-551-386 Sequence 386, App
c	414	7	1.0	440	9	US-10-180-998-386 Sequence 386, App
c	415	7	1.0	440	9	US-10-180-999-386 Sequence 386, App
c	416	7	1.0	440	9	US-10-183-013-386 Sequence 386, App
c	417	7	1.0	440	9	US-10-184-612-386 Sequence 386, App
c	418	7	1.0	440	9	US-10-184-616-386 Sequence 386, App
c	419	7	1.0	440	9	US-10-184-617-386 Sequence 386, App
c	420	7	1.0	440	9	US-10-184-622-386 Sequence 386, App
c	421	7	1.0	440	9	US-10-184-628-386 Sequence 386, App
c	422	7	1.0	440	9	US-10-184-629-386 Sequence 386, App
c	423	7	1.0	440	9	US-10-184-630-386 Sequence 386, App
c	424	7	1.0	440	9	US-10-184-631-386 Sequence 386, App
c	425	7	1.0	440	9	US-10-184-632-386 Sequence 386, App
c	426	7	1.0	440	9	US-10-184-636-386 Sequence 386, App
c	427	7	1.0	440	9	US-10-184-640-386 Sequence 386, App
c	428	7	1.0	440	9	US-10-184-650-386 Sequence 386, App
c	429	7	1.0	440	9	US-10-184-651-386 Sequence 386, App
c	430	7	1.0	440	9	US-10-187-588-386 Sequence 386, App
c	431	7	1.0	440	9	US-10-187-597-386 Sequence 386, App
c	432	7	1.0	440	9	US-10-187-598-386 Sequence 386, App
c	433	7	1.0	440	9	US-10-187-600-386 Sequence 386, App
c	434	7	1.0	440	9	US-10-187-601-386 Sequence 386, App
c	435	7	1.0	440	9	US-10-187-602-386 Sequence 386, App
c	436	7	1.0	440	9	US-10-187-603-386 Sequence 386, App
c	437	7	1.0	440	9	US-10-187-741-386 Sequence 386, App
c	438	7	1.0	440	9	US-10-187-743-386 Sequence 386, App
c	439	7	1.0	440	9	US-10-187-746-386 Sequence 386, App
c	440	7	1.0	440	9	US-10-187-747-386 Sequence 386, App
c	441	7	1.0	440	9	US-10-187-751-386 Sequence 386, App
c	442	7	1.0	440	9	US-10-187-753-386 Sequence 386, App
c	443	7	1.0	440	9	US-10-187-754-386 Sequence 386, App
c	444	7	1.0	440	9	US-10-187-757-386 Sequence 386, App
c	445	7	1.0	440	9	US-10-187-884-386 Sequence 386, App
c	446	7	1.0	440	9	US-10-188-767-386 Sequence 386, App
c	447	7	1.0	440	9	US-10-188-769-386 Sequence 386, App
c	448	7	1.0	440	9	US-10-188-770-386 Sequence 386, App
c	449	7	1.0	440	9	US-10-188-773-386 Sequence 386, App
c	450	7	1.0	440	9	US-10-188-781-386 Sequence 386, App
c	451	7	1.0	440	9	US-10-194-361-386 Sequence 386, App
c	452	7	1.0	440	9	US-10-194-423-386 Sequence 386, App
c	453	7	1.0	440	9	US-10-195-897-386 Sequence 386, App
c	454	7	1.0	440	9	US-10-195-901-386 Sequence 386, App
c	455	7	1.0	440	9	US-10-196-756-386 Sequence 386, App
c	456	7	1.0	440	9	US-10-173-708-386 Sequence 386, App
c	457	7	1.0	440	9	US-10-176-479-386 Sequence 386, App
c	458	7	1.0	440	9	US-10-176-748-386 Sequence 386, App
c	459	7	1.0	440	9	US-10-176-916-386 Sequence 386, App
c	460	7	1.0	440	9	US-10-179-507-386 Sequence 386, App
c	461	7	1.0	440	9	US-10-179-516-386 Sequence 386, App
c	462	7	1.0	440	9	US-10-179-519-386 Sequence 386, App
c	463	7	1.0	440	9	US-10-179-525-386 Sequence 386, App
c	464	7	1.0	440	9	US-10-180-540-386 Sequence 386, App
c	465	7	1.0	440	9	US-10-180-545-386 Sequence 386, App
c	466	7	1.0	440	9	US-10-183-006-386 Sequence 386, App
c	467	7	1.0	440	9	US-10-183-008-386 Sequence 386, App
c	468	7	1.0	440	9	US-10-183-017-386 Sequence 386, App
c	469	7	1.0	440	9	US-10-183-019-386 Sequence 386, App
c	470	7	1.0	440	9	US-10-184-618-386 Sequence 386, App
c	471	7	1.0	440	9	US-10-184-625-386 Sequence 386, App
c	472	7	1.0	440	9	US-10-184-626-386 Sequence 386, App
c	473	7	1.0	440	9	US-10-184-627-386 Sequence 386, App
c	474	7	1.0	440	9	US-10-184-645-386 Sequence 386, App
c	475	7	1.0	440	9	US-10-184-654-386 Sequence 386, App
c	476	7	1.0	440	9	US-10-184-655-386 Sequence 386, App
c	477	7	1.0	440	9	US-10-188-774-386 Sequence 386, App
c	478	7	1.0	440	9	US-10-188-775-386 Sequence 386, App
c	479	7	1.0	440	9	US-10-194-462-386 Sequence 386, App
c	480	7	1.0	440	9	US-10-195-902-386 Sequence 386, App
c	481	7	1.0	440	9	US-10-196-743-386 Sequence 386, App
c	482	7	1.0	440	9	US-10-196-745-386 Sequence 386, App
c	483	7	1.0	440	9	US-10-196-760-386 Sequence 386, App
c	484	7	1.0	440	9	US-10-196-762-386 Sequence 386, App
c	485	7	1.0	440	9	US-10-197-695-386 Sequence 386, App
c	486	7	1.0	440	9	US-10-176-484-386 Sequence 386, App
c	487	7	1.0	440	9	US-10-176-753-386 Sequence 386, App
c	488	7	1.0	440	9	US-10-176-917-386 Sequence 386, App
c	489	7	1.0	440	9	US-10-176-982-386 Sequence 386, App
c	490	7	1.0	440	9	US-10-179-506-386 Sequence 386, App
c	491	7	1.0	440	9	US-10-179-513-386 Sequence 386, App
c	492	7	1.0	440	9	US-10-179-514-386 Sequence 386, App
c	493	7	1.0	440	9	US-10-179-522-386 Sequence 386, App
c	494	7	1.0	440	9	US-10-180-556-386 Sequence 386, App
c	495	7	1.0	440	9	US-10-180-560-386 Sequence 386, App
c	496	7	1.0	440	9	US-10-183-015-386 Sequence 386, App
c	497	7	1.0	440	9	US-10-184-615-386 Sequence 386, App
c	498	7	1.0	440	9	US-10-184-620-386 Sequence 386, App
c	499	7	1.0	440	9	US-10-184-643-386 Sequence 386, App
c	500	7	1.0	440	9	US-10-184-656-386 Sequence 386, App
c	501	7	1.0	440	9	US-10-192-010-386 Sequence 386, App
c	502	7	1.0	440	9	US-10-195-894-386 Sequence 386, App
c	503	7	1.0	440	9	US-10-205-908-386 Sequence 386, App
c	504	7	1.0	440	9	US-10-063-518-134 Sequence 134, App
c	505	7	1.0	440	9	US-10-063-598-134 Sequence 134, App
c	506	7	1.0	440	9	US-10-184-619-386 Sequence 386, App
c	507	7	1.0	440	9	US-10-186-855-386 Sequence 386, App
c	508	7	1.0	440	9	US-10-187-599-386 Sequence 386, App
c	509	7	1.0	440	9	US-10-187-750-386 Sequence 386, App
c	510	7	1.0	440	9	US-10-188-780-386 Sequence 386, App
c	511	7	1.0	440	9	US-10-192-015-386 Sequence 386, App
c	512	7	1.0	440	9	US-10-194-394-386 Sequence 386, App
c	513	7	1.0	440	9	US-10-194-425-386 Sequence 386, App
c	514	7	1.0	440	9	US-10-194-485-386 Sequence 386, App
c	515	7	1.0	440	9	US-10-195-885-386 Sequence 386, App

c 516	7	1.0	440	9	US-10-195-890-386	Sequence 386, App
c 517	7	1.0	440	9	US-10-195-899-386	Sequence 386, App
c 518	7	1.0	440	9	US-10-196-748-386	Sequence 386, App
c 519	7	1.0	440	9	US-10-196-750-386	Sequence 386, App
c 520	7	1.0	440	9	US-10-197-699-386	Sequence 386, App
c 521	7	1.0	440	9	US-10-197-700-386	Sequence 386, App
c 522	7	1.0	440	9	US-10-197-705-386	Sequence 386, App
c 523	7	1.0	440	9	US-10-197-708-386	Sequence 386, App
c 524	7	1.0	440	9	US-10-198-764-386	Sequence 386, App
c 525	7	1.0	440	9	US-10-198-765-386	Sequence 386, App
c 526	7	1.0	440	9	US-10-198-768-386	Sequence 386, App
c 527	7	1.0	440	9	US-10-198-769-386	Sequence 386, App
c 528	7	1.0	440	9	US-10-199-305-386	Sequence 386, App
c 529	7	1.0	440	9	US-10-199-306-386	Sequence 386, App
c 530	7	1.0	440	9	US-10-199-310-386	Sequence 386, App
c 531	7	1.0	440	9	US-10-199-311-386	Sequence 386, App
c 532	7	1.0	440	9	US-10-199-314-386	Sequence 386, App
c 533	7	1.0	440	9	US-10-199-317-386	Sequence 386, App
c 534	7	1.0	440	9	US-10-199-665-386	Sequence 386, App
c 535	7	1.0	440	9	US-10-199-666-386	Sequence 386, App
c 536	7	1.0	440	9	US-10-199-669-386	Sequence 386, App
c 537	7	1.0	440	9	US-10-201-534-386	Sequence 386, App
c 538	7	1.0	440	9	US-10-201-770-386	Sequence 386, App
c 539	7	1.0	440	9	US-10-201-855-386	Sequence 386, App
c 540	7	1.0	440	9	US-10-201-856-386	Sequence 386, App
c 541	7	1.0	440	9	US-10-202-469-386	Sequence 386, App
c 542	7	1.0	440	9	US-10-202-470-386	Sequence 386, App
c 543	7	1.0	440	9	US-10-202-476-386	Sequence 386, App
c 544	7	1.0	440	9	US-10-202-934-386	Sequence 386, App
c 545	7	1.0	440	9	US-10-202-935-386	Sequence 386, App
c 546	7	1.0	440	9	US-10-202-936-386	Sequence 386, App
c 547	7	1.0	440	9	US-10-202-939-386	Sequence 386, App
c 548	7	1.0	440	9	US-10-205-504-386	Sequence 386, App
c 549	7	1.0	440	9	US-10-205-509-386	Sequence 386, App
c 550	7	1.0	440	9	US-10-205-895-386	Sequence 386, App
c 551	7	1.0	440	9	US-10-205-899-386	Sequence 386, App
c 552	7	1.0	440	9	US-10-205-900-386	Sequence 386, App
c 553	7	1.0	440	9	US-10-205-909-386	Sequence 386, App
c 554	7	1.0	440	9	US-10-227-693-134	Sequence 134, App
c 555	7	1.0	440	9	US-10-183-002-386	Sequence 386, App
c 556	7	1.0	440	9	US-10-184-621-386	Sequence 386, App
c 557	7	1.0	440	9	US-10-184-638-386	Sequence 386, App
c 558	7	1.0	440	9	US-10-187-752-386	Sequence 386, App
c 559	7	1.0	440	9	US-10-187-887-386	Sequence 386, App
c 560	7	1.0	440	9	US-10-194-461-386	Sequence 386, App
c 561	7	1.0	440	9	US-10-195-892-386	Sequence 386, App
c 562	7	1.0	440	9	US-10-196-751-386	Sequence 386, App
c 563	7	1.0	440	9	US-10-197-694-386	Sequence 386, App
c 564	7	1.0	440	9	US-10-197-697-386	Sequence 386, App
c 565	7	1.0	440	9	US-10-197-707-386	Sequence 386, App
c 566	7	1.0	440	9	US-10-199-303-386	Sequence 386, App
c 567	7	1.0	440	9	US-10-199-318-386	Sequence 386, App
c 568	7	1.0	440	9	US-10-199-458-386	Sequence 386, App
c 569	7	1.0	440	9	US-10-199-462-386	Sequence 386, App
c 570	7	1.0	440	9	US-10-201-324-386	Sequence 386, App
c 571	7	1.0	440	9	US-10-201-328-386	Sequence 386, App
c 572	7	1.0	440	9	US-10-201-527-386	Sequence 386, App
c 573	7	1.0	440	9	US-10-201-528-386	Sequence 386, App
c 574	7	1.0	440	9	US-10-201-529-386	Sequence 386, App
c 575	7	1.0	440	9	US-10-201-530-386	Sequence 386, App
c 576	7	1.0	440	9	US-10-202-408-386	Sequence 386, App
c 577	7	1.0	440	9	US-10-202-409-386	Sequence 386, App
c 578	7	1.0	440	9	US-10-202-411-386	Sequence 386, App
c 579	7	1.0	440	9	US-10-202-472-386	Sequence 386, App
c 580	7	1.0	440	9	US-10-205-502-386	Sequence 386, App
c 581	7	1.0	440	9	US-10-205-507-386	Sequence 386, App
c 582	7	1.0	440	9	US-10-205-511-386	Sequence 386, App
c 583	7	1.0	440	9	US-10-205-902-386	Sequence 386, App
c 584	7	1.0	440	9	US-10-205-907-386	Sequence 386, App
c 585	7	1.0	440	9	US-10-194-456-386	Sequence 386, App
c 586	7	1.0	440	9	US-10-196-758-386	Sequence 386, App
c 587	7	1.0	440	9	US-10-198-770-386	Sequence 386, App
c 588	7	1.0	440	9	US-10-199-308-386	Sequence 386, App
c 589	7	1.0	440	9	US-10-200-617-386	Sequence 386, App
c 590	7	1.0	440	9	US-10-205-893-386	Sequence 386, App
c 591	7	1.0	440	9	US-10-205-897-386	Sequence 386, App
c 592	7	1.0	440	9	US-10-063-567-134	Sequence 134, App
c 593	7	1.0	440	9	US-10-174-571-386	Sequence 386, App
c 594	7	1.0	440	9	US-10-176-746-386	Sequence 386, App
c 595	7	1.0	440	9	US-10-176-923-386	Sequence 386, App
c 596	7	1.0	440	9	US-10-183-011-386	Sequence 386, App
c 597	7	1.0	440	9	US-10-184-633-386	Sequence 386, App
c 598	7	1.0	440	9	US-10-184-639-386	Sequence 386, App
c 599	7	1.0	440	9	US-10-187-742-386	Sequence 386, App
c 600	7	1.0	440	9	US-10-187-748-386	Sequence 386, App
c 601	7	1.0	440	9	US-10-188-766-386	Sequence 386, App
c 602	7	1.0	440	9	US-10-188-771-386	Sequence 386, App
c 603	7	1.0	440	9	US-10-192-006-386	Sequence 386, App
c 604	7	1.0	440	9	US-10-192-008-386	Sequence 386, App
c 605	7	1.0	440	9	US-10-192-009-386	Sequence 386, App
c 606	7	1.0	440	9	US-10-192-012-386	Sequence 386, App
c 607	7	1.0	440	9	US-10-192-014-386	Sequence 386, App
c 608	7	1.0	440	9	US-10-192-016-386	Sequence 386, App
c 609	7	1.0	440	9	US-10-194-362-386	Sequence 386, App
c 610	7	1.0	440	9	US-10-194-364-386	Sequence 386, App
c 611	7	1.0	440	9	US-10-194-395-386	Sequence 386, App
c 612	7	1.0	440	9	US-10-194-424-386	Sequence 386, App
c 613	7	1.0	440	9	US-10-194-458-386	Sequence 386, App
c 614	7	1.0	440	9	US-10-194-459-386	Sequence 386, App
c 615	7	1.0	440	9	US-10-194-488-386	Sequence 386, App
c 616	7	1.0	440	9	US-10-195-886-386	Sequence 386, App
c 617	7	1.0	440	9	US-10-195-891-386	Sequence 386, App
c 618	7	1.0	440	9	US-10-196-746-386	Sequence 386, App
c 619	7	1.0	440	9	US-10-196-752-386	Sequence 386, App
c 620	7	1.0	440	9	US-10-196-753-386	Sequence 386, App
c 621	7	1.0	440	9	US-10-196-754-386	Sequence 386, App
c 622	7	1.0	440	9	US-10-196-761-386	Sequence 386, App
c 623	7	1.0	440	9	US-10-197-692-386	Sequence 386, App
c 624	7	1.0	440	9	US-10-197-693-386	Sequence 386, App
c 625	7	1.0	440	9	US-10-197-696-386	Sequence 386, App
c 626	7	1.0	440	9	US-10-197-698-386	Sequence 386, App
c 627	7	1.0	440	9	US-10-197-703-386	Sequence 386, App
c 628	7	1.0	440	9	US-10-197-711-386	Sequence 386, App
c 629	7	1.0	440	9	US-10-198-757-386	Sequence 386, App
c 630	7	1.0	440	9	US-10-198-761-386	Sequence 386, App
c 631	7	1.0	440	9	US-10-198-762-386	Sequence 386, App
c 632	7	1.0	440	9	US-10-198-763-386	Sequence 386, App
c 633	7	1.0	440	9	US-10-198-767-386	Sequence 386, App
c 634	7	1.0	440	9	US-10-199-301-386	Sequence 386, App
c 635	7	1.0	440	9	US-10-199-307-386	Sequence 386, App
c 636	7	1.0	440	9	US-10-199-312-386	Sequence 386, App
c 637	7	1.0	440	9	US-10-199-315-386	Sequence 386, App
c 638	7	1.0	440	9	US-10-199-316-386	Sequence 386, App
c 639	7	1.0	440	9	US-10-199-457-386	Sequence 386, App
c 640	7	1.0	440	9	US-10-199-459-386	Sequence 386, App
c 641	7	1.0	440	9	US-10-199-460-386	Sequence 386, App
c 642	7	1.0	440	9	US-10-199-461-386	Sequence 386, App
c 643	7	1.0	440	9	US-10-199-667-386	Sequence 386, App
c 644	7	1.0	440	9	US-10-199-673-386	Sequence 386, App
c 645	7	1.0	440	9	US-10-201-321-386	Sequence 386, App
c 646	7	1.0	440	9	US-10-201-322-386	Sequence 386, App
c 647	7	1.0	440	9	US-10-201-326-386	Sequence 386, App
c 648	7	1.0	440	9	US-10-201-532-386	Sequence 386, App
c 649	7	1.0	440	9	US-10-201-533-386	Sequence 386, App
c 650	7	1.0	440	9	US-10-201-535-386	Sequence 386, App
c 651	7	1.0	440	9	US-10-201-769-386	Sequence 386, App
c 652	7	1.0	440	9	US-10-201-771-386	Sequence 386, App
c 653	7	1.0	440	9	US-10-201-854-386	Sequence 386, App
c 654	7	1.0	440	9	US-10-202-410-386	Sequence 386, App
c 655	7	1.0	440	9	US-10-202-473-386	Sequence 386, App
c 656	7	1.0	440	9	US-10-202-474-386	Sequence 386, App
c 657	7	1.0	440	9	US-10-205-503-386	Sequence 386, App
c 658	7	1.0	440	9	US-10-205-512-386	Sequence 386, App
c 659	7	1.0	440	9	US-10-205-892-386	Sequence 386, App
c 660	7	1.0	440	9	US-10-205-894-386	Sequence 386, App
c 661	7	1.0	440	9	US-10-205-896-386	Sequence 386, App

c	662	7	1.0	440	9	US-10-205-898-386 Sequence 386, App
c	663	7	1.0	440	9	US-10-205-901-386 Sequence 386, App
c	664	7	1.0	440	9	US-10-205-903-386 Sequence 386, App
c	665	7	1.0	440	9	US-10-206-909-386 Sequence 386, App
c	666	7	1.0	440	9	US-10-206-910-386 Sequence 386, App
c	667	7	1.0	440	9	US-10-206-911-386 Sequence 386, App
c	668	7	1.0	440	9	US-10-206-912-386 Sequence 386, App
c	669	7	1.0	440	9	US-10-206-913-386 Sequence 386, App
c	670	7	1.0	440	9	US-10-206-914-386 Sequence 386, App
c	671	7	1.0	440	9	US-10-206-920-386 Sequence 386, App
c	672	7	1.0	440	9	US-10-206-921-386 Sequence 386, App
c	673	7	1.0	440	9	US-10-206-923-386 Sequence 386, App
c	674	7	1.0	440	9	US-10-206-925-386 Sequence 386, App
c	675	7	1.0	440	9	US-10-206-926-386 Sequence 386, App
c	676	7	1.0	440	9	US-10-206-927-386 Sequence 386, App
c	677	7	1.0	440	9	US-10-207-916-386 Sequence 386, App
c	678	7	1.0	440	9	US-10-207-917-386 Sequence 386, App
c	679	7	1.0	440	9	US-10-207-918-386 Sequence 386, App
c	680	7	1.0	440	9	US-10-207-919-386 Sequence 386, App
c	681	7	1.0	440	9	US-10-207-920-386 Sequence 386, App
c	682	7	1.0	440	9	US-10-207-925-386 Sequence 386, App
c	683	7	1.0	440	9	US-10-208-021-386 Sequence 386, App
c	684	7	1.0	440	9	US-10-208-022-386 Sequence 386, App
c	685	7	1.0	440	9	US-10-208-023-386 Sequence 386, App
c	686	7	1.0	440	9	US-10-208-026-386 Sequence 386, App
c	687	7	1.0	440	9	US-10-208-029-386 Sequence 386, App
c	688	7	1.0	440	9	US-10-208-030-386 Sequence 386, App
c	689	7	1.0	440	9	US-10-232-232-386 Sequence 386, App
c	690	7	1.0	440	9	US-10-063-538-134 Sequence 134, App
c	691	7	1.0	440	9	US-10-173-693-386 Sequence 386, App
c	692	7	1.0	440	9	US-10-174-578-386 Sequence 386, App
c	693	7	1.0	440	9	US-10-175-741-386 Sequence 386, App
c	694	7	1.0	440	9	US-10-175-750-386 Sequence 386, App
c	695	7	1.0	440	9	US-10-176-986-386 Sequence 386, App
c	696	7	1.0	440	9	US-10-184-641-386 Sequence 386, App
c	697	7	1.0	440	9	US-10-187-888-386 Sequence 386, App
c	698	7	1.0	440	9	US-10-194-360-386 Sequence 386, App
c	699	7	1.0	440	9	US-10-194-365-386 Sequence 386, App
c	700	7	1.0	440	9	US-10-195-895-386 Sequence 386, App
c	701	7	1.0	440	9	US-10-195-898-386 Sequence 386, App
c	702	7	1.0	440	9	US-10-196-759-386 Sequence 386, App
c	703	7	1.0	440	9	US-10-199-302-386 Sequence 386, App
c	704	7	1.0	440	9	US-10-201-323-386 Sequence 386, App
c	705	7	1.0	440	9	US-10-205-510-386 Sequence 386, App
c	706	7	1.0	440	9	US-10-205-891-386 Sequence 386, App
c	707	7	1.0	440	9	US-10-205-904-386 Sequence 386, App
c	708	7	1.0	440	9	US-10-206-917-386 Sequence 386, App
c	709	7	1.0	440	9	US-10-207-923-386 Sequence 386, App
c	710	7	1.0	440	9	US-10-207-924-386 Sequence 386, App
c	711	7	1.0	440	9	US-10-208-028-386 Sequence 386, App
c	712	7	1.0	440	9	US-10-063-599-134 Sequence 134, App
c	713	7	1.0	440	9	US-10-121-062-386 Sequence 386, App
c	714	7	1.0	440	9	US-10-175-753-386 Sequence 386, App
c	715	7	1.0	440	9	US-10-180-553-386 Sequence 386, App
c	716	7	1.0	440	9	US-10-201-327-386 Sequence 386, App
c	717	7	1.0	440	9	US-10-063-595-134 Sequence 134, App
c	718	7	1.0	440	9	US-10-173-696-386 Sequence 386, App
c	719	7	1.0	440	9	US-10-183-003-386 Sequence 386, App
c	720	7	1.0	440	9	US-10-183-016-386 Sequence 386, App
c	721	7	1.0	440	9	US-10-125-923A-386 Sequence 386, App
c	722	7	1.0	440	9	US-10-176-491-386 Sequence 386, App
c	723	7	1.0	440	9	US-10-176-979-386 Sequence 386, App
c	724	7	1.0	440	9	US-10-187-592-386 Sequence 386, App
c	725	7	1.0	440	9	US-10-197-691-386 Sequence 386, App
c	726	7	1.0	440	9	US-10-198-771-386 Sequence 386, App
c	727	7	1.0	440	9	US-10-174-575A-386 Sequence 386, App
c	728	7	1.0	440	9	US-10-179-520-386 Sequence 386, App
c	729	7	1.0	440	9	US-10-201-325-386 Sequence 386, App
c	730	7	1.0	440	9	US-10-202-941-386 Sequence 386, App
c	731	7	1.0	440	9	US-10-205-910-386 Sequence 386, App
c	732	7	1.0	440	9	US-10-179-526-386 Sequence 386, App
c	733	7	1.0	440	12	US-10-006-867-134 Sequence 134, App
c	734	7	1.0	440	12	US-10-052-586-386 Sequence 386, App

	735	7	1.0	448	10	US-09-815-242-12574 Sequence 12574, A
c	736	7	1.0	449	9	US-09-990-046-29 Sequence 29, Appl
c	737	7	1.0	449	9	US-09-733-634-30 Sequence 30, Appl
	738	7	1.0	449	10	US-09-764-864-1163 Sequence 1163, Ap
	739	7	1.0	452	10	US-09-815-242-10493 Sequence 10493, A
c	740	7	1.0	453	10	US-09-815-242-10872 Sequence 10872, A
	741	7	1.0	454	10	US-09-729-674-140 Sequence 140, App
c	742	7	1.0	456	9	US-09-764-868-985 Sequence 985, App
	743	7	1.0	458	10	US-09-739-457-4 Sequence 4, Appli
	744	7	1.0	459	9	US-09-769-787-85 Sequence 85, Appl
	745	7	1.0	460	9	US-09-738-626-6145 Sequence 6145, Ap
	746	7	1.0	464	9	US-09-988-462-22 Sequence 22, Appl
	747	7	1.0	469	10	US-09-925-302-627 Sequence 627, App
c	748	7	1.0	480	9	US-10-084-018-8 Sequence 8, Appli
c	749	7	1.0	480	9	US-09-736-457-336 Sequence 336, App
c	750	7	1.0	480	9	US-09-902-941-336 Sequence 336, App
c	751	7	1.0	480	9	US-09-849-626-336 Sequence 336, App
c	752	7	1.0	480	9	US-10-017-754-336 Sequence 336, App
c	753	7	1.0	481	10	US-09-764-864-979 Sequence 979, App
c	754	7	1.0	482	9	US-09-746-783-66 Sequence 66, Appl
	755	7	1.0	483	10	US-09-799-777-69 Sequence 69, Appl
c	756	7	1.0	484	10	US-09-815-242-10809 Sequence 10809, A
	757	7	1.0	486	9	US-10-055-475-12 Sequence 12, Appl
	758	7	1.0	490	10	US-09-739-457-3 Sequence 3, Appli
	759	7	1.0	490	10	US-09-739-457-6 Sequence 6, Appli
	760	7	1.0	490	10	US-09-739-457-7 Sequence 7, Appli
	761	7	1.0	492	9	US-09-852-797-72 Sequence 72, Appl
	762	7	1.0	492	10	US-09-853-161-72 Sequence 72, Appl
	763	7	1.0	492	10	US-09-852-659A-72 Sequence 72, Appl
	764	7	1.0	501	9	US-10-051-909-38 Sequence 38, Appl
c	765	7	1.0	504	10	US-09-883-797-6 Sequence 6, Appli
c	766	7	1.0	513	10	US-09-818-264-4 Sequence 4, Appli
c	767	7	1.0	518	9	US-09-976-297-2 Sequence 2, Appli
c	768	7	1.0	519	10	US-09-925-300-1680 Sequence 1680, Ap
c	769	7	1.0	542	9	US-10-045-815-8 Sequence 8, Appli
	770	7	1.0	549	9	US-10-055-475-16 Sequence 16, Appl
	771	7	1.0	549	10	US-09-828-313-39 Sequence 39, Appl
c	772	7	1.0	550	9	US-10-095-139-16 Sequence 16, Appl
c	773	7	1.0	550	9	US-10-339-513-2 Sequence 2, Appli
c	774	7	1.0	553	9	US-09-292-862-2 Sequence 2, Appli
c	775	7	1.0	553	9	US-10-233-873A-7 Sequence 7, Appli
c	776	7	1.0	553	10	US-09-746-359A-11 Sequence 11, Appl
c	777	7	1.0	553	10	US-09-949-192-7 Sequence 7, Appli
	778	7	1.0	563	10	US-09-764-864-1070 Sequence 1070, Ap
c	779	7	1.0	571	10	US-09-764-864-1028 Sequence 1028, Ap
	780	7	1.0	579	10	US-09-815-242-5821 Sequence 5821, Ap
	781	7	1.0	579	10	US-09-815-242-12943 Sequence 12943, A
	782	7	1.0	584	9	US-10-050-786-7 Sequence 7, Appli
c	783	7	1.0	585	9	US-10-173-519-4 Sequence 4, Appli
	784	7	1.0	587	9	US-10-173-519-5 Sequence 5, Appli
	785	7	1.0	587	9	US-10-167-994-12 Sequence 12, Appl
	786	7	1.0	594	9	US-10-028-072-10 Sequence 10, Appl
	787	7	1.0	594	9	US-10-121-049-10 Sequence 10, Appl
	788	7	1.0	594	9	US-10-123-904-10 Sequence 10, Appl
	789	7	1.0	594	9	US-10-140-470-10 Sequence 10, Appl
	790	7	1.0	594	9	US-10-175-746-10 Sequence 10, Appl
	791	7	1.0	594	9	US-10-176-918-10 Sequence 10, Appl
	792	7	1.0	594	9	US-10-176-921-10 Sequence 10, Appl
	793	7	1.0	594	9	US-10-137-865-10 Sequence 10, Appl
	794	7	1.0	594	9	US-10-140-474-10 Sequence 10, Appl
	795	7	1.0	594	9	US-10-142-431-10 Sequence 10, Appl
	796	7	1.0	594	9	US-10-143-114-10 Sequence 10, Appl
	797	7	1.0	594	9	US-10-140-002-10 Sequence 10, Appl
	798	7	1.0	594	9	US-10-142-419-10 Sequence 10, Appl
	799	7	1.0	594	9	US-10-123-262-10 Sequence 10, Appl
	800	7	1.0	594	9	US-10-142-423-10 Sequence 10, Appl
	801	7	1.0	594	9	US-10-121-050-10 Sequence 10, Appl
	802	7	1.0	594	9	US-10-141-755-10 Sequence 10, Appl
	803	7	1.0	594	9	US-10-143-032-10 Sequence 10, Appl
	804	7	1.0	594	9	US-10-123-108-10 Sequence 10, Appl
	805	7	1.0	594	9	US-10-123-236-10 Sequence 10, Appl
	806	7	1.0	594	9	US-10-123-261-10 Sequence 10, Appl
	807	7	1.0	594	9	US-10-140-921-10 Sequence 10, Appl

	808	7	1.0	594	9	US-10-140-928-10	Sequence 10, Appl
	809	7	1.0	594	9	US-10-121-045-10	Sequence 10, Appl
	810	7	1.0	594	9	US-10-123-292-10	Sequence 10, Appl
	811	7	1.0	594	9	US-10-123-903-10	Sequence 10, Appl
	812	7	1.0	594	9	US-10-124-819-10	Sequence 10, Appl
	813	7	1.0	594	9	US-10-124-822-10	Sequence 10, Appl
	814	7	1.0	594	9	US-10-140-925-10	Sequence 10, Appl
	815	7	1.0	594	9	US-10-160-498-10	Sequence 10, Appl
	816	7	1.0	594	9	US-10-121-041-10	Sequence 10, Appl
	817	7	1.0	594	9	US-10-121-043-10	Sequence 10, Appl
	818	7	1.0	594	9	US-10-121-047-10	Sequence 10, Appl
	819	7	1.0	594	9	US-10-123-215-10	Sequence 10, Appl
	820	7	1.0	594	9	US-10-123-902-10	Sequence 10, Appl
	821	7	1.0	594	9	US-10-123-908-10	Sequence 10, Appl
	822	7	1.0	594	9	US-10-123-909-10	Sequence 10, Appl
	823	7	1.0	594	9	US-10-123-910-10	Sequence 10, Appl
	824	7	1.0	594	9	US-10-124-813-10	Sequence 10, Appl
	825	7	1.0	594	9	US-10-124-817-10	Sequence 10, Appl
	826	7	1.0	594	9	US-10-124-824-10	Sequence 10, Appl
	827	7	1.0	594	9	US-10-125-922-10	Sequence 10, Appl
	828	7	1.0	594	9	US-10-125-924-10	Sequence 10, Appl
	829	7	1.0	594	9	US-10-127-825A-10	Sequence 10, Appl
	830	7	1.0	594	9	US-10-127-829A-10	Sequence 10, Appl
	831	7	1.0	594	9	US-10-127-835A-10	Sequence 10, Appl
	832	7	1.0	594	9	US-10-127-839A-10	Sequence 10, Appl
	833	7	1.0	594	9	US-10-127-901A-10	Sequence 10, Appl
	834	7	1.0	594	9	US-10-128-693A-10	Sequence 10, Appl
	835	7	1.0	594	9	US-10-131-813A-10	Sequence 10, Appl
	836	7	1.0	594	9	US-10-131-818A-10	Sequence 10, Appl
	837	7	1.0	594	9	US-10-131-823A-10	Sequence 10, Appl
	838	7	1.0	594	9	US-10-131-824A-10	Sequence 10, Appl
	839	7	1.0	594	9	US-10-131-830A-10	Sequence 10, Appl
	840	7	1.0	594	9	US-10-131-837A-10	Sequence 10, Appl
	841	7	1.0	594	9	US-10-137-872A-10	Sequence 10, Appl
	842	7	1.0	594	9	US-10-140-860-10	Sequence 10, Appl
	843	7	1.0	594	9	US-10-142-417-10	Sequence 10, Appl
	844	7	1.0	594	9	US-10-147-500-10	Sequence 10, Appl
	845	7	1.0	594	9	US-10-147-502-10	Sequence 10, Appl
	846	7	1.0	594	9	US-10-147-515-10	Sequence 10, Appl
	847	7	1.0	594	9	US-10-147-517-10	Sequence 10, Appl
	848	7	1.0	594	9	US-10-147-519-10	Sequence 10, Appl
	849	7	1.0	594	9	US-10-147-526-10	Sequence 10, Appl
	850	7	1.0	594	9	US-10-147-527-10	Sequence 10, Appl
	851	7	1.0	594	9	US-10-152-395-10	Sequence 10, Appl
	852	7	1.0	594	9	US-10-157-782-10	Sequence 10, Appl
	853	7	1.0	594	9	US-10-121-040-10	Sequence 10, Appl
	854	7	1.0	594	9	US-10-121-056-10	Sequence 10, Appl
	855	7	1.0	594	9	US-10-121-061-10	Sequence 10, Appl
	856	7	1.0	594	9	US-10-123-235-10	Sequence 10, Appl
	857	7	1.0	594	9	US-10-124-818-10	Sequence 10, Appl
	858	7	1.0	594	9	US-10-125-926A-10	Sequence 10, Appl
	859	7	1.0	594	9	US-10-125-930A-10	Sequence 10, Appl
	860	7	1.0	594	9	US-10-127-831A-10	Sequence 10, Appl
	861	7	1.0	594	9	US-10-127-837A-10	Sequence 10, Appl
	862	7	1.0	594	9	US-10-127-838B-10	Sequence 10, Appl
	863	7	1.0	594	9	US-10-127-842A-10	Sequence 10, Appl
	864	7	1.0	594	9	US-10-127-843A-10	Sequence 10, Appl
	865	7	1.0	594	9	US-10-127-845A-10	Sequence 10, Appl
	866	7	1.0	594	9	US-10-127-846A-10	Sequence 10, Appl
	867	7	1.0	594	9	US-10-127-848A-10	Sequence 10, Appl
	868	7	1.0	594	9	US-10-127-849A-10	Sequence 10, Appl
	869	7	1.0	594	9	US-10-127-850A-10	Sequence 10, Appl
	870	7	1.0	594	9	US-10-127-851A-10	Sequence 10, Appl
	871	7	1.0	594	9	US-10-128-684A-10	Sequence 10, Appl
	872	7	1.0	594	9	US-10-128-686A-10	Sequence 10, Appl
	873	7	1.0	594	9	US-10-128-690A-10	Sequence 10, Appl
	874	7	1.0	594	9	US-10-128-691A-10	Sequence 10, Appl
	875	7	1.0	594	9	US-10-131-819A-10	Sequence 10, Appl
	876	7	1.0	594	9	US-10-131-829A-10	Sequence 10, Appl
	877	7	1.0	594	9	US-10-131-836A-10	Sequence 10, Appl
	878	7	1.0	594	9	US-10-137-868-10	Sequence 10, Appl
	879	7	1.0	594	9	US-10-146-729-10	Sequence 10, Appl
	880	7	1.0	594	9	US-10-146-791-10	Sequence 10, Appl
	881	7	1.0	594	9	US-10-147-484-10	Sequence 10, Appl
	882	7	1.0	594	9	US-10-147-492-10	Sequence 10, Appl
	883	7	1.0	594	9	US-10-147-508-10	Sequence 10, Appl
	884	7	1.0	594	9	US-10-147-512-10	Sequence 10, Appl
	885	7	1.0	594	9	US-10-158-782-10	Sequence 10, Appl
	886	7	1.0	594	9	US-10-175-735-10	Sequence 10, Appl
	887	7	1.0	594	9	US-10-123-905-10	Sequence 10, Appl
	888	7	1.0	594	9	US-10-123-907-10	Sequence 10, Appl
	889	7	1.0	594	9	US-10-124-815-10	Sequence 10, Appl
	890	7	1.0	594	9	US-10-125-921A-10	Sequence 10, Appl
	891	7	1.0	594	9	US-10-125-928A-10	Sequence 10, Appl
	892	7	1.0	594	9	US-10-127-821A-10	Sequence 10, Appl
	893	7	1.0	594	9	US-10-127-822A-10	Sequence 10, Appl
	894	7	1.0	594	9	US-10-127-824A-10	Sequence 10, Appl
	895	7	1.0	594	9	US-10-127-826A-10	Sequence 10, Appl
	896	7	1.0	594	9	US-10-127-827A-10	Sequence 10, Appl
	897	7	1.0	594	9	US-10-127-828A-10	Sequence 10, Appl
	898	7	1.0	594	9	US-10-127-830A-10	Sequence 10, Appl
	899	7	1.0	594	9	US-10-127-832A-10	Sequence 10, Appl
	900	7	1.0	594	9	US-10-127-833A-10	Sequence 10, Appl
	901	7	1.0	594	9	US-10-127-834A-10	Sequence 10, Appl
	902	7	1.0	594	9	US-10-127-836A-10	Sequence 10, Appl
	903	7	1.0	594	9	US-10-127-841A-10	Sequence 10, Appl
	904	7	1.0	594	9	US-10-127-844A-10	Sequence 10, Appl
	905	7	1.0	594	9	US-10-128-687A-10	Sequence 10, Appl
	906	7	1.0	594	9	US-10-128-688A-10	Sequence 10, Appl
	907	7	1.0	594	9	US-10-128-689A-10	Sequence 10, Appl
	908	7	1.0	594	9	US-10-128-694A-10	Sequence 10, Appl
	909	7	1.0	594	9	US-10-131-825A-10	Sequence 10, Appl
	910	7	1.0	594	9	US-10-230-417-10	Sequence 10, Appl
	911	7	1.0	594	9	US-10-121-051-10	Sequence 10, Appl
	912	7	1.0	594	9	US-10-131-815A-10	Sequence 10, Appl
	913	7	1.0	594	9	US-10-131-817A-10	Sequence 10, Appl
	914	7	1.0	594	9	US-10-131-821A-10	Sequence 10, Appl
	915	7	1.0	594	9	US-10-131-822A-10	Sequence 10, Appl
	916	7	1.0	594	9	US-10-131-828A-10	Sequence 10, Appl
	917	7	1.0	594	9	US-10-131-835A-10	Sequence 10, Appl
	918	7	1.0	594	9	US-10-137-864A-10	Sequence 10, Appl
	919	7	1.0	594	9	US-10-137-869A-10	Sequence 10, Appl
	920	7	1.0	594	9	US-10-147-523-10	Sequence 10, Appl
	921	7	1.0	594	9	US-10-158-785-10	Sequence 10, Appl
	922	7	1.0	594	9	US-10-121-042-10	Sequence 10, Appl
	923	7	1.0	594	9	US-10-123-912-10	Sequence 10, Appl
c	924	7	1.0	612	9	US-09-875-353-4	Sequence 4, Appli
	925	7	1.0	623	9	US-09-815-379-2	Sequence 2, Appli
c	926	7	1.0	633	9	US-10-086-464-11	Sequence 11, Appl
	927	7	1.0	635	9	US-10-214-932-86	Sequence 86, Appl
	928	7	1.0	637	9	US-09-948-137A-4	Sequence 4, Appli
c	929	7	1.0	652	10	US-09-789-919-96	Sequence 96, Appl
c	930	7	1.0	653	9	US-10-023-437-67	Sequence 67, Appl
c	931	7	1.0	653	9	US-09-820-843A-26	Sequence 26, Appl
	932	7	1.0	695	9	US-09-738-626-5527	Sequence 5527, Ap
c	933	7	1.0	695	10	US-09-810-796-15	Sequence 15, Appl
c	934	7	1.0	696	10	US-09-813-148-6	Sequence 6, Appli
	935	7	1.0	704	10	US-09-815-242-11805	Sequence 11805, A
	936	7	1.0	705	9	US-10-043-487-269	Sequence 269, App
	937	7	1.0	707	9	US-10-068-674-25	Sequence 25, Appl
c	938	7	1.0	722	9	US-10-128-870-23	Sequence 23, Appl
c	939	7	1.0	722	9	US-10-131-685-23	Sequence 23, Appl
	940	7	1.0	724	9	US-10-108-605-307	Sequence 307, App
c	941	7	1.0	747	9	US-10-045-815-6	Sequence 6, Appli
c	942	7	1.0	763	9	US-09-989-339-4	Sequence 4, Appli
c	943	7	1.0	763	9	US-09-769-787-63	Sequence 63, Appl
	944	7	1.0	765	12	US-10-044-303-1	Sequence 1, Appli
	945	7	1.0	786	10	US-09-815-242-10103	Sequence 10103, A
c	946	7	1.0	808	9	US-10-108-605-127	Sequence 127, App
	947	7	1.0	815	10	US-09-371-307-88	Sequence 88, Appl
	948	7	1.0	816	9	US-09-796-720B-2	Sequence 2, Appli
	949	7	1.0	819	9	US-09-796-720B-4	Sequence 4, Appli
	950	7	1.0	823	10	US-09-815-242-13348	Sequence 13348, A
	951	7	1.0	840	10	US-09-815-242-5198	Sequence 5198, Ap
c	952	7	1.0	844	10	US-09-813-148-4	Sequence 4, Appli
c	953	7	1.0	844	10	US-09-810-796-14	Sequence 14, Appl

c 954	7	1.0	871	9	US-10-128-870-20	Sequence 20, Appl
c 955	7	1.0	871	9	US-10-131-685-20	Sequence 20, Appl
956	7	1.0	878	9	US-10-169-048-42	Sequence 42, Appl
957	7	1.0	878	10	US-09-912-020-364	Sequence 364, App
c 958	7	1.0	881	10	US-09-982-736-2	Sequence 2, Appli
c 959	7	1.0	888	10	US-09-810-796-5	Sequence 5, Appli
c 960	7	1.0	890	9	US-09-373-658-4	Sequence 4, Appli
c 961	7	1.0	897	10	US-09-810-796-4	Sequence 4, Appli
c 962	7	1.0	911	12	US-10-141-132-2	Sequence 2, Appli
c 963	7	1.0	922	9	US-09-908-153B-42	Sequence 42, Appl
c 964	7	1.0	923	10	US-09-825-147-2	Sequence 2, Appli
c 965	7	1.0	926	10	US-09-816-685-2	Sequence 2, Appli
c 966	7	1.0	932	10	US-09-813-148-2	Sequence 2, Appli
967	7	1.0	966	10	US-09-771-161A-197	Sequence 197, App
968	7	1.0	982	9	US-09-991-496-95	Sequence 95, Appl
c 969	7	1.0	982	10	US-09-841-132-176	Sequence 176, App
970	7	1.0	982	10	US-09-874-923-95	Sequence 95, Appl
971	7	1.0	1002	9	US-09-988-117-3	Sequence 3, Appli
972	7	1.0	1002	10	US-09-812-471-3	Sequence 3, Appli
973	7	1.0	1002	10	US-09-812-633-3	Sequence 3, Appli
c 974	7	1.0	1006	10	US-09-841-132-190	Sequence 190, App
c 975	7	1.0	1014	9	US-10-054-044A-4	Sequence 4, Appli
c 976	7	1.0	1014	9	US-10-120-604-4	Sequence 4, Appli
c 977	7	1.0	1014	9	US-10-120-604-139	Sequence 139, App
978	7	1.0	1016	12	US-10-044-303-2	Sequence 2, Appli
c 979	7	1.0	1021	9	US-10-101-464A-954	Sequence 954, App
c 980	7	1.0	1164	10	US-09-950-046A-2	Sequence 2, Appli
981	7	1.0	1242	9	US-09-925-299-911	Sequence 911, App
982	7	1.0	1242	10	US-09-925-299-911	Sequence 911, App
983	7	1.0	1300	10	US-09-815-242-4903	Sequence 4903, Ap
984	7	1.0	1300	10	US-09-815-242-10906	Sequence 10906, A
c 985	7	1.0	1306	9	US-10-017-724-4	Sequence 4, Appli
c 986	7	1.0	1311	1	US-08-954-701A-4	Sequence 4, Appli
987	7	1.0	1338	9	US-10-029-413A-10	Sequence 10, Appl
988	7	1.0	1356	9	US-10-090-280-33	Sequence 33, Appl
989	7	1.0	1359	9	US-10-090-280-34	Sequence 34, Appl
c 990	7	1.0	1375	10	US-09-740-274-4	Sequence 4, Appli
c 991	7	1.0	1400	9	US-10-123-036-4	Sequence 4, Appli
992	7	1.0	1403	9	US-10-108-605-93	Sequence 93, Appl
993	7	1.0	1518	10	US-09-801-368-152	Sequence 152, App
c 994	7	1.0	1620	10	US-09-827-949-2	Sequence 2, Appli
995	7	1.0	1641	9	US-09-991-496-96	Sequence 96, Appl
996	7	1.0	1641	10	US-09-874-923-96	Sequence 96, Appl
c 997	7	1.0	1643	9	US-09-738-626-6894	Sequence 6894, Ap
998	7	1.0	1647	9	US-09-824-574-4	Sequence 4, Appli
c 999	7	1.0	1663	9	US-09-842-758-41	Sequence 41, Appl
c1000	7	1.0	1663	10	US-09-875-519A-22	Sequence 22, Appl

ALIGNMENTS

```
RESULT 1
US-09-764-891-3857
; Sequence 3857, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3857
;   LENGTH: 106
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (10)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-3857

  Alignment Scores:
  Pred. No.:              3.15e-68       Length:        106
  Score:                  84.00          Matches:       84
  Percent Similarity:     100.00%        Conservative:  0
  Best Local Similarity:  100.00%        Mismatches:    0
  Query Match:            11.60%         Indels:        0
  DB:                     9              Gaps:          0

US-09-714-865-15 (1-2172) x US-09-764-891-3857 (1-106)

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCG 1944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 GlyArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGluSer 30

Qy       1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 AspAsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValPro 50

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGA 2064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 AlaTrpLeuGluGluIleAlaPheSerThrTyrIleProGlyPheSerGlySerThrArg 70

Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGG 2124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GlyAsnValPheAlaSerValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGly 90

Qy       2125 TTTTCTTCTTCA 2136
              ||||||||||||
Db         91 PheSerSerSer 94

RESULT 2
US-10-195-117-3
; Sequence 3, Application US/10195117
; Publication No. US20030092083A1
; GENERAL INFORMATION:
;  APPLICANT: In2Gen Co., Ltd.
;  APPLICANT:  Jeoung, Doo-il
;  APPLICANT:  Cho, Bomsoo
;  APPLICANT:  Lim, Yoon
;  APPLICANT:  Park, Saeyoung
;  APPLICANT:  Lee, Daeyeon
;  APPLICANT:  Bang, Yung-Jue
;  APPLICANT:  Yang, Hankwang
;  APPLICANT:  Kim, Dae-Kee
;  TITLE OF INVENTION: CAGE Antigen
;  FILE REFERENCE: 59258-00002
;  CURRENT APPLICATION NUMBER: US/10/195,117
;  CURRENT FILING DATE:  2002-07-11
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-195-117-3

  Alignment Scores:
  Pred. No.:              6.54e-05       Length:        238
  Score:                  15.00          Matches:       15
  Percent Similarity:     100.00%        Conservative:  0
  Best Local Similarity:  100.00%        Mismatches:    0
  Query Match:            2.07%          Indels:        0
  DB:                     9              Gaps:          0

US-09-714-865-15 (1-2172) x US-10-195-117-3 (1-238)

Qy       1324 TACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              |||||||||||||||||||||||||||||||||||||||||||||
Db        153 TyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 167

RESULT 3
US-10-195-117-4
```

```
; Sequence 4, Application US/10195117
; Publication No. US20030092083A1
; GENERAL INFORMATION:
;  APPLICANT: In2Gen Co., Ltd.
;  APPLICANT:  Jeoung, Doo-il
;  APPLICANT:  Cho, Bomsoo
;  APPLICANT:  Lim, Yoon
;  APPLICANT:  Park, Saeyoung
;  APPLICANT:  Lee, Daeyeon
;  APPLICANT:  Bang, Yung-Jue
;  APPLICANT:  Yang, Hankwang
;  APPLICANT:  Kim, Dae-Kee
;  TITLE OF INVENTION: CAGE Antigen
;  FILE REFERENCE: 59258-00002
;  CURRENT APPLICATION NUMBER: US/10/195,117
;  CURRENT FILING DATE:  2002-07-11
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-195-117-4

Alignment Scores:
Pred. No.:                6.54e-05       Length:        238
Score:                    15.00          Matches:       15
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              2.07%          Indels:        0
DB:                       9              Gaps:          0

US-09-714-865-15 (1-2172) x US-10-195-117-4 (1-238)

Qy       1324 TACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              |||||||||||||||||||||||||||||||||||||||||||||
Db        153 TyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 167

RESULT 4
US-09-925-301-1439
; Sequence 1439, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1439
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (244)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (305)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (325)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (328)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (340)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1439

Alignment Scores:
Pred. No.:                6.17e-05       Length:        343
Score:                    15.00          Matches:       15
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              2.07%          Indels:        0
DB:                       10             Gaps:          0

US-09-714-865-15 (1-2172) x US-09-925-301-1439 (1-343)

Qy       1324 TACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              |||||||||||||||||||||||||||||||||||||||||||||
Db         18 TyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 32

RESULT 5
US-10-228-897-13
; Sequence 13, Application US/10228897
; Publication No. US20030092043A1
; GENERAL INFORMATION:
;  APPLICANT: Fisher, Paul B.
;  APPLICANT:  Kang, Dong-Chul
;  APPLICANT:  Gopalkrishnan, Rahul V.
;  TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
;  TITLE OF INVENTION:  GENE-5 AND PROMOTER AND USES THEREOF
;  FILE REFERENCE: A34614-A-PCT-USA (070050.2121)
;  CURRENT APPLICATION NUMBER: US/10/228,897
;  CURRENT FILING DATE:  2002-08-26
;  PRIOR APPLICATION NUMBER: PCT/US01/06960
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 09/515,363
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 500
;   TYPE: PRT
;   ORGANISM: molgula oculata
US-10-228-897-13

Alignment Scores:
Pred. No.:                5.81e-05       Length:        500
Score:                    15.00          Matches:       15
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              2.07%          Indels:        0
DB:                       9              Gaps:          0

US-09-714-865-15 (1-2172) x US-10-228-897-13 (1-500)

Qy       1324 TACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              |||||||||||||||||||||||||||||||||||||||||||||
Db        244 TyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 258

RESULT 6
US-10-108-605-55
; Sequence 55, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
;  APPLICANT: Broadus, Julie
;  APPLICANT:  Stam, Lynn
;  APPLICANT:  Bachmann, Jane
;  APPLICANT:  Kamdar, Kim
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
;  TITLE OF INVENTION:  PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
;  FILE REFERENCE: 31133B
;  CURRENT APPLICATION NUMBER: US/10/108,605
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: US 09/761,142
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/176,418
```

```
;  PRIOR FILING DATE: 2000-01-14
;  NUMBER OF SEQ ID NOS: 361
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
;   LENGTH: 575
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-108-605-55

Alignment Scores:
Pred. No.:              5.69e-05      Length:        575
Score:                  15.00         Matches:       15
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            2.07%         Indels:        0
DB:                     9             Gaps:          0

US-09-714-865-15 (1-2172) x US-10-108-605-55 (1-575)

Qy       1324 TACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              |||||||||||||||||||||||||||||||||||||||||||||
Db        287 TyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPhe 301

RESULT 7
US-09-864-761-46256
; Sequence 46256, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46256
;   LENGTH: 45
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC019205.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;   OTHER INFORMATION: EST_HUMAN HIT: AI637624.1, EVALUE 1.00e-19
;   OTHER INFORMATION: SWISSPROT HIT: P24782, EVALUE 9.00e-14
US-09-864-761-46256

Alignment Scores:
Pred. No.:              0.402         Length:        45
Score:                  11.00         Matches:       11
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            1.52%         Indels:        0
DB:                     10            Gaps:          0

US-09-714-865-15 (1-2172) x US-09-864-761-46256 (1-45)

Qy       1867 GAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
              |||||||||||||||||||||||||||||||||
Db         34 GluTyrValHisArgIleGlyArgThrGlyArg 44

RESULT 8
US-09-815-242-10990
; Sequence 10990, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10990
;   LENGTH: 613
;   TYPE: PRT
;   ORGANISM: Haemophilus influenzae
US-09-815-242-10990

Alignment Scores:
Pred. No.:              0.266         Length:        613
Score:                  11.00         Matches:       11
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
```

```
Query Match:             1.52%          Indels:        0
DB:                      10             Gaps:          0

US-09-714-865-15 (1-2172) x US-09-815-242-10990 (1-613)

Qy         994 GCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
               |||||||||||||||||||||||||||||||||
Db          49 AlaGlnThrGlySerGlyLysThrAlaAlaPhe 59

RESULT 9
US-09-815-242-10311
; Sequence 10311, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10311
;   LENGTH: 646
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-815-242-10311

Alignment Scores:
Pred. No.:               0.263          Length:        646
Score:                   11.00          Matches:       11
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             1.52%          Indels:        0
DB:                      10             Gaps:          0

US-09-714-865-15 (1-2172) x US-09-815-242-10311 (1-646)

Qy         994 GCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
               |||||||||||||||||||||||||||||||||
Db          67 AlaGlnThrGlySerGlyLysThrAlaAlaPhe 77

RESULT 10
US-09-815-242-13769
; Sequence 13769, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13769
;   LENGTH: 646
;   TYPE: PRT
;   ORGANISM: Salmonella typhi
US-09-815-242-13769

Alignment Scores:
Pred. No.:               0.263          Length:        646
Score:                   11.00          Matches:       11
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             1.52%          Indels:        0
DB:                      10             Gaps:          0

US-09-714-865-15 (1-2172) x US-09-815-242-13769 (1-646)

Qy         994 GCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
               |||||||||||||||||||||||||||||||||
Db          67 AlaGlnThrGlySerGlyLysThrAlaAlaPhe 77

RESULT 11
US-09-923-831-43
; Sequence 43, Application US/09923831
; Patent No. US20020115142A1
; GENERAL INFORMATION:
;  APPLICANT: Martelange, Val,rie
;  APPLICANT:  De Smet, Charles
;  APPLICANT:  Boon-Falleur, Thierry
;  TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE: L0461/7054
;  CURRENT APPLICATION NUMBER: US/09/923,831
;  CURRENT FILING DATE:  2001-08-07
;  PRIOR APPLICATION NUMBER: 09/183,706
;  PRIOR FILING DATE: 2001-10-30
;  NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 43
;   LENGTH: 648
;   TYPE: PRT
;   ORGANISM: H. sapiens
US-09-923-831-43

Alignment Scores:
Pred. No.:               0.263          Length:        648
Score:                   11.00          Matches:       11
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             1.52%          Indels:        0
DB:                      10             Gaps:          0
```

```
US-09-714-865-15 (1-2172) x US-09-923-831-43 (1-648)

Qy       1867 GAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
               |||||||||||||||||||||||||||||||||
Db        570 GluTyrValHisArgIleGlyArgThrGlyArg 580

RESULT 12
US-09-815-242-11703
; Sequence 11703, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11703
;   LENGTH: 421
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-815-242-11703

Alignment Scores:
Pred. No.:               2.34        Length:        421
Score:                   10.00       Matches:       10
Percent Similarity:      100.00%     Conservative:  0
Best Local Similarity:   100.00%     Mismatches:    0
Query Match:             1.38%       Indels:        0
DB:                      10          Gaps:          0

US-09-714-865-15 (1-2172) x US-09-815-242-11703 (1-421)

Qy       1870 TATGTTCATCGAATTGGGCGTACTGGTCGT 1899
               ||||||||||||||||||||||||||||||
Db        340 TyrValHisArgIleGlyArgThrGlyArg 349

RESULT 13
US-09-738-626-4359
; Sequence 4359, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4359
;   LENGTH: 424
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-4359

Alignment Scores:
Pred. No.:               2.33        Length:        424
Score:                   10.00       Matches:       10
Percent Similarity:      100.00%     Conservative:  0
Best Local Similarity:   100.00%     Mismatches:    0
Query Match:             1.38%       Indels:        0
DB:                      9           Gaps:          0

US-09-714-865-15 (1-2172) x US-09-738-626-4359 (1-424)

Qy       1870 TATGTTCATCGAATTGGGCGTACTGGTCGT 1899
               ||||||||||||||||||||||||||||||
Db        338 TyrValHisArgIleGlyArgThrGlyArg 347

RESULT 14
US-09-971-536-52
; Sequence 52, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
;  APPLICANT: Glenn, Matthew
;  APPLICANT:  Havukkala, Ilkka
;  APPLICANT:  Bloksberg, Leonard
;  APPLICANT:  Lubbers, Mark
;  APPLICANT:  Dekker, James
;  APPLICANT:  Christensson, Anna
;  APPLICANT:  Holland, Ross
;  APPLICANT:  O'Toole, Paul
;  APPLICANT:  Reid, Julian
;  APPLICANT:  Coolbear, Timothy
;  TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Methods
;  TITLE OF INVENTION:  Using Them
;  FILE REFERENCE: 1043c2
;  CURRENT APPLICATION NUMBER: US/09/971,536
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
;  PRIOR FILING DATE: 2000-11-28
;  PRIOR APPLICATION NUMBER: PCT/NZ01/00160
;  PRIOR FILING DATE: 2001-08-08
;  NUMBER OF SEQ ID NOS: 83
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52
;   LENGTH: 495
;   TYPE: PRT
;   ORGANISM: Lactobacillus rhamnosus
US-09-971-536-52

Alignment Scores:
Pred. No.:               2.28        Length:        495
Score:                   10.00       Matches:       10
Percent Similarity:      100.00%     Conservative:  0
Best Local Similarity:   100.00%     Mismatches:    0
```

```
Query Match:              1.38%          Indels:        0
DB:                       9              Gaps:          0

US-09-714-865-15 (1-2172) x US-09-971-536-52 (1-495)

Qy        1870 TATGTTCATCGAATTGGGCGTACTGGTCGT 1899
               ||||||||||||||||||||||||||||||
Db         316 TyrValHisArgIleGlyArgThrGlyArg 325

RESULT 15
US-09-815-242-13491
; Sequence 13491, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13491
;    LENGTH: 524
;    TYPE: PRT
;    ORGANISM: Streptococcus pneumoniae
US-09-815-242-13491

Alignment Scores:
Pred. No.:                2.26           Length:        524
Score:                    10.00          Matches:       10
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              1.38%          Indels:        0
DB:                       10             Gaps:          0

US-09-714-865-15 (1-2172) x US-09-815-242-13491 (1-524)

Qy        1870 TATGTTCATCGAATTGGGCGTACTGGTCGT 1899
               ||||||||||||||||||||||||||||||
Db         323 TyrValHisArgIleGlyArgThrGlyArg 332


Search completed: June 10, 2003, 17:22:48
Job time : 53.5 secs
```

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:          June 10, 2003, 17:05:01 ; Search time 18.5 Seconds
                                          (without alignments)
                                          6908.820 Million cell updates/sec

Title:           US-09-714-865-15
Perfect score:   724
Sequence:        1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:   OLIGO
                 Xgapop 60.0 , Xgapext 60.0
                 Ygapop 60.0 , Ygapext 60.0
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        262574 seqs, 29422922 residues

Word size:       1

Total number of hits satisfying chosen parameters:      438908

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111834_26626/app_query.fasta_1.2311
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=oligo.rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=1000 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_16_@runat_05062003_111834_26626 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :       Issued_Patents_AA:*
                 1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                 2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                 3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                 4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                 5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                 6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	3.0	662	3	US-09-058-489-15	Sequence 15, Appl
	2	22	3.0	662	3	US-09-058-489-16	Sequence 16, Appl
	3	18	2.5	660	3	US-09-058-489-18	Sequence 18, Appl
	4	18	2.5	660	3	US-09-058-489-91	Sequence 91, Appl
	5	11	1.5	648	4	US-09-183-706-43	Sequence 43, Appl
	6	11	1.5	648	4	US-09-567-995-43	Sequence 43, Appl
	7	10	1.4	149	4	US-08-679-493A-85	Sequence 85, Appl
c	8	9	1.3	191	4	US-09-393-627B-25	Sequence 25, Appl
c	9	9	1.3	543	2	US-08-224-482-4	Sequence 4, Appli
	10	9	1.2	746	4	US-09-149-934-4	Sequence 4, Appli
c	11	9	1.3	870	2	US-09-010-928B-2	Sequence 2, Appli
	12	9	1.2	1261	4	US-09-208-742-4	Sequence 4, Appli
	13	9	1.2	1261	4	US-09-332-295-2	Sequence 2, Appli
	14	9	1.2	1261	4	US-09-709-979-2	Sequence 2, Appli
c	15	8	1.1	17	1	US-08-264-093-24	Sequence 24, Appl
	16	8	1.1	78	2	US-08-245-511-20	Sequence 20, Appl
	17	8	1.1	78	2	US-08-600-993A-20	Sequence 20, Appl
c	18	8	1.1	86	4	US-09-009-816-8	Sequence 8, Appli
c	19	8	1.1	112	2	US-08-678-194-6	Sequence 6, Appli
c	20	8	1.1	112	4	US-08-890-011-6	Sequence 6, Appli
c	21	8	1.1	112	4	US-09-262-724-6	Sequence 6, Appli
c	22	8	1.1	113	1	US-08-264-093-10	Sequence 10, Appl
	23	8	1.1	166	4	US-09-134-001C-4279	Sequence 4279, Ap
c	24	8	1.1	191	4	US-09-393-627B-23	Sequence 23, Appl
c	25	8	1.1	271	2	US-08-467-265-2	Sequence 2, Appli
c	26	8	1.1	271	4	US-08-467-265-2	Sequence 2, Appli
c	27	8	1.1	271	4	US-09-407-891-2	Sequence 2, Appli
c	28	8	1.1	367	4	US-09-009-816-2	Sequence 2, Appli
	29	8	1.1	479	4	US-09-149-934-3	Sequence 3, Appli
	30	8	1.1	528	4	US-09-134-001C-4262	Sequence 4262, Ap
c	31	8	1.1	535	4	US-08-983-035A-38	Sequence 38, Appl
	32	8	1.1	547	3	US-09-176-657-1	Sequence 1, Appli
	33	8	1.1	752	2	US-08-896-590A-2	Sequence 2, Appli
	34	8	1.1	752	2	US-08-896-590A-4	Sequence 4, Appli
	35	8	1.1	752	4	US-09-185-832-2	Sequence 2, Appli
	36	8	1.1	752	4	US-09-185-832-4	Sequence 4, Appli
	37	8	1.1	859	4	US-09-149-934-1	Sequence 1, Appli
c	38	7	1.0	19	1	US-07-908-317-7	Sequence 7, Appli
	39	7	1.0	19	4	US-08-630-915A-170	Sequence 170, App
c	40	7	1.0	19	5	PCT-US93-06171-7	Sequence 7, Appli
	41	7	1.0	22	4	US-09-177-249-302	Sequence 302, App
	42	7	1.0	23	4	US-09-098-901-7	Sequence 7, Appli
	43	7	1.0	23	4	US-09-098-901-11	Sequence 11, Appl
	44	7	1.0	41	2	US-07-921-447-10	Sequence 10, Appl
c	45	7	1.0	48	2	US-08-370-156-9	Sequence 9, Appli
c	46	7	1.0	48	3	US-08-990-065-22	Sequence 22, Appl
c	47	7	1.0	48	4	US-08-975-084-4	Sequence 4, Appli
	48	7	1.0	53	4	US-09-288-143-81	Sequence 81, Appl
c	49	7	1.0	56	6	5496550-10	Patent No. 5496550
c	50	7	1.0	61	4	US-09-227-357-621	Sequence 621, App
c	51	7	1.0	64	4	US-09-461-697-267	Sequence 267, App
c	52	7	1.0	68	2	US-08-484-434C-21	Sequence 21, Appl
c	53	7	1.0	75	4	US-09-461-697-265	Sequence 265, App
	54	7	1.0	76	4	US-09-605-785-575	Sequence 575, App
	55	7	1.0	78	2	US-08-245-511-39	Sequence 39, Appl
	56	7	1.0	78	2	US-08-600-993A-39	Sequence 39, Appl
c	57	7	1.0	85	4	US-09-430-564-8	Sequence 8, Appli
c	58	7	1.0	85	4	US-09-430-564-14	Sequence 14, Appl
	59	7	1.0	86	2	US-08-245-511-38	Sequence 38, Appl
	60	7	1.0	86	2	US-08-600-993A-38	Sequence 38, Appl
	61	7	1.0	88	4	US-08-936-165A-513	Sequence 513, App
c	62	7	1.0	91	1	US-08-469-202-18	Sequence 18, Appl
	63	7	1.0	109	4	US-09-134-001C-3265	Sequence 3265, Ap
c	64	7	1.0	128	4	US-09-134-001C-2881	Sequence 2881, Ap
c	65	7	1.0	130	6	5268359-2	Patent No. 5268359
c	66	7	1.0	136	2	US-08-665-647-15	Sequence 15, Appl
c	67	7	1.0	137	1	US-08-392-828C-38	Sequence 38, Appl
c	68	7	1.0	137	3	US-09-330-945-38	Sequence 38, Appl
c	69	7	1.0	144	1	US-08-133-979A-21	Sequence 21, Appl
c	70	7	1.0	144	2	US-08-436-890-21	Sequence 21, Appl
c	71	7	1.0	144	2	US-08-451-213-21	Sequence 21, Appl
c	72	7	1.0	155	1	US-08-530-010-14	Sequence 14, Appl
c	73	7	1.0	155	2	US-08-484-101B-14	Sequence 14, Appl
c	74	7	1.0	155	4	US-08-714-524D-14	Sequence 14, Appl
c	75	7	1.0	170	4	US-08-259-451-19	Sequence 19, Appl
c	76	7	1.0	171	5	PCT-US95-04910-12	Sequence 12, Appl
c	77	7	1.0	176	4	US-09-411-722-1	Sequence 1, Appli
	78	7	1.0	177	4	US-09-318-443-4	Sequence 4, Appli
c	79	7	1.0	188	4	US-09-724-864-48	Sequence 48, Appl
c	80	7	1.0	192	1	US-08-086-428B-62	Sequence 62, Appl
c	81	7	1.0	192	1	US-08-086-428B-64	Sequence 64, Appl
c	82	7	1.0	192	2	US-08-468-570-62	Sequence 62, Appl
c	83	7	1.0	192	2	US-08-468-570-64	Sequence 64, Appl
c	84	7	1.0	192	2	US-08-290-665A-62	Sequence 62, Appl
c	85	7	1.0	192	2	US-08-290-665A-64	Sequence 64, Appl

c	86	7	1.0	192	5	PCT-US95-10398-62	Sequence 62, Appl
c	87	7	1.0	192	5	PCT-US95-10398-64	Sequence 64, Appl
c	88	7	1.0	199	3	US-08-506-553C-8	Sequence 8, Appli
	89	7	1.0	205	4	US-09-637-481-2	Sequence 2, Appli
	90	7	1.0	205	4	US-09-723-428-2	Sequence 2, Appli
c	91	7	1.0	208	2	US-08-606-143-41	Sequence 41, Appl
c	92	7	1.0	210	4	US-08-612-973-14	Sequence 14, Appl
c	93	7	1.0	210	4	US-08-927-597-14	Sequence 14, Appl
c	94	7	1.0	216	3	US-08-506-553C-23	Sequence 23, Appl
c	95	7	1.0	224	2	US-08-616-857-2	Sequence 2, Appli
c	96	7	1.0	228	4	US-09-724-864-42	Sequence 42, Appl
c	97	7	1.0	244	1	US-08-289-699A-3	Sequence 3, Appli
c	98	7	1.0	244	2	US-08-878-283-3	Sequence 3, Appli
c	99	7	1.0	244	4	US-09-182-616-3	Sequence 3, Appli
	100	7	1.0	245	4	US-09-202-548B-9	Sequence 9, Appli
c	101	7	1.0	245	4	US-09-105-058C-8	Sequence 8, Appli
	102	7	1.0	247	3	US-09-167-434-3	Sequence 3, Appli
	103	7	1.0	247	3	US-09-167-434-4	Sequence 4, Appli
	104	7	1.0	247	3	US-08-853-755-3	Sequence 3, Appli
	105	7	1.0	247	3	US-08-853-755-4	Sequence 4, Appli
c	106	7	1.0	247	5	PCT-US94-10257A-2	Sequence 2, Appli
	107	7	1.0	257	1	US-08-451-947-10	Sequence 10, Appl
	108	7	1.0	257	2	US-08-424-826A-10	Sequence 10, Appl
	109	7	1.0	257	3	US-08-928-694-10	Sequence 10, Appl
	110	7	1.0	257	5	PCT-US91-06950-10	Sequence 10, Appl
	111	7	1.0	259	3	US-09-082-090-2	Sequence 2, Appli
c	112	7	1.0	268	1	US-08-431-387-4	Sequence 4, Appli
	113	7	1.0	274	4	US-09-723-430-2	Sequence 2, Appli
	114	7	1.0	274	4	US-09-632-155-2	Sequence 2, Appli
	115	7	1.0	274	4	US-09-724-518-2	Sequence 2, Appli
c	116	7	1.0	276	3	US-08-506-553C-26	Sequence 26, Appl
c	117	7	1.0	281	1	US-08-040-548-3	Sequence 3, Appli
c	118	7	1.0	281	1	US-08-466-344-3	Sequence 3, Appli
c	119	7	1.0	300	3	US-08-765-856-2	Sequence 2, Appli
c	120	7	1.0	300	4	US-08-935-009A-2	Sequence 2, Appli
c	121	7	1.0	300	4	US-09-105-058C-4	Sequence 4, Appli
c	122	7	1.0	300	4	US-09-105-058C-6	Sequence 6, Appli
c	123	7	1.0	301	3	US-08-829-525-24	Sequence 24, Appl
c	124	7	1.0	301	4	US-08-609-583A-24	Sequence 24, Appl
c	125	7	1.0	301	4	US-08-937-399-24	Sequence 24, Appl
c	126	7	1.0	301	4	US-09-310-367-24	Sequence 24, Appl
c	127	7	1.0	301	4	US-09-032-337-24	Sequence 24, Appl
c	128	7	1.0	302	3	US-08-765-856-4	Sequence 4, Appli
c	129	7	1.0	302	4	US-08-935-009A-4	Sequence 4, Appli
c	130	7	1.0	305	4	US-09-134-001C-5632	Sequence 5632, Ap
	131	7	1.0	306	1	US-08-279-754-2	Sequence 2, Appli
	132	7	1.0	306	5	PCT-US95-09052-2	Sequence 2, Appli
	133	7	1.0	317	1	US-08-221-750A-13	Sequence 13, Appl
c	134	7	1.0	326	2	US-09-055-097-6	Sequence 6, Appli
	135	7	1.0	335	4	US-08-736-915-2	Sequence 2, Appli
	136	7	1.0	337	4	US-09-641-806-4	Sequence 4, Appli
	137	7	1.0	341	4	US-09-724-517-4	Sequence 4, Appli
	138	7	1.0	341	4	US-09-641-807A-4	Sequence 4, Appli
	139	7	1.0	341	4	US-09-723-096-4	Sequence 4, Appli
	140	7	1.0	342	4	US-09-641-806-2	Sequence 2, Appli
c	141	7	1.0	347	3	US-08-445-515-58	Sequence 58, Appl
c	142	7	1.0	348	3	US-08-445-515-56	Sequence 56, Appl
c	143	7	1.0	349	3	US-09-009-620-2	Sequence 2, Appli
	144	7	1.0	350	1	US-08-458-067-2	Sequence 2, Appli
	145	7	1.0	350	5	PCT-US96-07795-2	Sequence 2, Appli
	146	7	1.0	350	5	PCT-US96-07796-2	Sequence 2, Appli
	147	7	1.0	357	4	US-09-718-841-4	Sequence 4, Appli
	148	7	1.0	357	4	US-09-718-810-4	Sequence 4, Appli
c	149	7	1.0	359	4	US-09-444-336-1	Sequence 1, Appli
c	150	7	1.0	361	3	US-09-034-985-2	Sequence 2, Appli
	151	7	1.0	361	4	US-09-596-541-2	Sequence 2, Appli
	152	7	1.0	362	4	US-09-594-669-6	Sequence 6, Appli
c	153	7	1.0	362	4	US-09-444-336-2	Sequence 2, Appli
	154	7	1.0	369	4	US-09-596-541-6	Sequence 6, Appli
	155	7	1.0	370	4	US-09-596-541-4	Sequence 4, Appli
	156	7	1.0	375	4	US-09-572-191-4	Sequence 4, Appli
	157	7	1.0	375	4	US-09-723-262-4	Sequence 4, Appli
	158	7	1.0	375	4	US-09-723-219-4	Sequence 4, Appli
c	159	7	1.0	378	2	US-09-055-097-1	Sequence 1, Appli
	160	7	1.0	378	4	US-09-186-489-2	Sequence 2, Appli
	161	7	1.0	387	4	US-09-175-928-2	Sequence 2, Appli
	162	7	1.0	391	4	US-09-594-669-4	Sequence 4, Appli
c	163	7	1.0	393	2	US-09-026-587-3	Sequence 3, Appli
c	164	7	1.0	393	2	US-09-227-420-3	Sequence 3, Appli
c	165	7	1.0	397	1	US-08-647-928-8	Sequence 8, Appli
	166	7	1.0	401	2	US-08-549-004A-5	Sequence 5, Appli
c	167	7	1.0	401	3	US-08-517-802-3	Sequence 3, Appli
	168	7	1.0	401	4	US-09-051-982A-5	Sequence 5, Appli
	169	7	1.0	403	2	US-08-533-669A-10	Sequence 10, Appl
	170	7	1.0	403	2	US-08-607-509-2	Sequence 2, Appli
	171	7	1.0	403	2	US-08-607-509-4	Sequence 4, Appli
	172	7	1.0	403	2	US-08-454-036-2	Sequence 2, Appli
	173	7	1.0	403	2	US-08-634-642-2	Sequence 2, Appli
	174	7	1.0	403	2	US-08-634-642-4	Sequence 4, Appli
	175	7	1.0	403	3	US-08-989-370-2	Sequence 2, Appli
	176	7	1.0	403	3	US-08-989-370-4	Sequence 4, Appli
	177	7	1.0	403	4	US-09-183-861-10	Sequence 10, Appl
	178	7	1.0	403	4	US-09-022-765-10	Sequence 10, Appl
	179	7	1.0	403	5	PCT-US95-05064-2	Sequence 2, Appli
c	180	7	1.0	404	1	US-07-602-824A-4	Sequence 4, Appli
c	181	7	1.0	404	1	US-07-602-608-4	Sequence 4, Appli
c	182	7	1.0	404	1	US-07-983-451-4	Sequence 4, Appli
c	183	7	1.0	404	1	US-08-261-578-4	Sequence 4, Appli
c	184	7	1.0	404	1	US-08-261-577-10	Sequence 10, Appl
	185	7	1.0	405	4	US-09-594-669-12	Sequence 12, Appl
	186	7	1.0	408	1	US-07-951-715A-21	Sequence 21, Appl
	187	7	1.0	408	2	US-08-459-448A-21	Sequence 21, Appl
	188	7	1.0	408	3	US-08-459-595A-21	Sequence 21, Appl
	189	7	1.0	408	3	US-08-459-504B-21	Sequence 21, Appl
	190	7	1.0	408	3	US-08-459-444-21	Sequence 21, Appl
	191	7	1.0	408	4	US-09-547-422-21	Sequence 21, Appl
	192	7	1.0	409	4	US-09-572-191-6	Sequence 6, Appli
	193	7	1.0	409	4	US-09-723-262-6	Sequence 6, Appli
	194	7	1.0	409	4	US-09-723-219-6	Sequence 6, Appli
c	195	7	1.0	411	4	US-08-460-900C-10	Sequence 10, Appl
	196	7	1.0	411	4	US-09-318-443-6	Sequence 6, Appli
	197	7	1.0	411	4	US-09-318-443-8	Sequence 8, Appli
c	198	7	1.0	411	4	US-08-954-698-10	Sequence 10, Appl
c	199	7	1.0	411	4	US-08-957-874-10	Sequence 10, Appl
c	200	7	1.0	411	4	US-09-325-256-19	Sequence 19, Appl
c	201	7	1.0	414	5	PCT-US92-06840-2	Sequence 2, Appli
	202	7	1.0	415	4	US-09-318-443-2	Sequence 2, Appli
c	203	7	1.0	418	2	US-09-026-587-1	Sequence 1, Appli
c	204	7	1.0	418	2	US-09-227-420-1	Sequence 1, Appli
c	205	7	1.0	430	2	US-08-924-440-2	Sequence 2, Appli
	206	7	1.0	430	4	US-09-594-669-2	Sequence 2, Appli
	207	7	1.0	434	4	US-09-594-669-10	Sequence 10, Appl
	208	7	1.0	448	2	US-08-959-749-2	Sequence 2, Appli
	209	7	1.0	448	4	US-09-351-497-2	Sequence 2, Appli
c	210	7	1.0	449	4	US-09-293-505-29	Sequence 29, Appl
	211	7	1.0	455	4	US-09-134-001C-3653	Sequence 3653, Ap
	212	7	1.0	464	1	US-07-951-715A-22	Sequence 22, Appl
	213	7	1.0	464	2	US-08-459-448A-22	Sequence 22, Appl
	214	7	1.0	464	3	US-08-459-595A-22	Sequence 22, Appl
	215	7	1.0	464	3	US-08-459-504B-22	Sequence 22, Appl
	216	7	1.0	464	3	US-08-459-444-22	Sequence 0, Appli
	217	7	1.0	464	4	US-09-547-422-22	Sequence 0, Appli
c	218	7	1.0	467	4	US-09-126-420A-17	Sequence 17, Appl
c	219	7	1.0	469	2	US-08-477-451-23	Sequence 23, Appl
	220	7	1.0	473	4	US-09-594-669-8	Sequence 8, Appli
	221	7	1.0	473	4	US-09-592-054-6	Sequence 6, Appli
c	222	7	1.0	480	2	US-08-828-488-8	Sequence 8, Appli
c	223	7	1.0	480	3	US-09-189-035-5	Sequence 5, Appli
c	224	7	1.0	480	4	US-09-382-086-5	Sequence 5, Appli
c	225	7	1.0	480	4	US-09-299-689A-8	Sequence 8, Appli
c	226	7	1.0	482	4	US-09-430-564-2	Sequence 2, Appli
	227	7	1.0	483	3	US-09-027-166-7	Sequence 7, Appli
c	228	7	1.0	485	1	US-08-453-956-15	Sequence 15, Appl
c	229	7	1.0	485	1	US-08-086-631-15	Sequence 15, Appl
c	230	7	1.0	485	2	US-08-452-930-15	Sequence 15, Appl
c	231	7	1.0	485	5	PCT-US93-08174-15	Sequence 15, Appl

c 232	7	1.0	488	2	US-08-928-692-10	Sequence 10, Appl
c 233	7	1.0	488	4	US-09-339-972-10	Sequence 10, Appl
c 234	7	1.0	492	1	US-07-794-393-4	Sequence 4, Appli
c 235	7	1.0	492	1	US-08-001-711-4	Sequence 4, Appli
236	7	1.0	492	4	US-09-152-060-72	Sequence 72, Appl
c 237	7	1.0	495	2	US-08-749-902-3	Sequence 3, Appli
c 238	7	1.0	495	4	US-09-430-564-3	Sequence 3, Appli
c 239	7	1.0	496	2	US-08-224-482-2	Sequence 2, Appli
c 240	7	1.0	504	4	US-08-868-373-6	Sequence 6, Appli
c 241	7	1.0	517	3	US-09-189-760-2	Sequence 2, Appli
c 242	7	1.0	517	4	US-09-514-422-2	Sequence 2, Appli
c 243	7	1.0	518	3	US-09-043-123-2	Sequence 2, Appli
244	7	1.0	522	4	US-09-592-054-4	Sequence 4, Appli
c 245	7	1.0	533	1	US-08-040-548-1	Sequence 1, Appli
c 246	7	1.0	533	1	US-08-466-344-1	Sequence 1, Appli
c 247	7	1.0	533	6	5206152-2	Patent No. 5206152
c 248	7	1.0	534	2	US-08-356-786-10	Sequence 10, Appl
249	7	1.0	534	4	US-09-103-664A-2	Sequence 2, Appli
250	7	1.0	544	4	US-08-687-590-30	Sequence 30, Appl
c 251	7	1.0	552	3	US-09-295-186-10	Sequence 10, Appl
c 252	7	1.0	553	2	US-08-943-087-2	Sequence 2, Appli
c 253	7	1.0	553	2	US-08-943-087-14	Sequence 14, Appl
c 254	7	1.0	553	2	US-08-943-087-16	Sequence 16, Appl
c 255	7	1.0	553	2	US-08-943-087-18	Sequence 18, Appl
c 256	7	1.0	553	2	US-08-943-087-20	Sequence 20, Appl
c 257	7	1.0	553	2	US-08-943-087-22	Sequence 22, Appl
c 258	7	1.0	553	2	US-08-943-087-24	Sequence 24, Appl
c 259	7	1.0	553	2	US-08-943-087-26	Sequence 26, Appl
c 260	7	1.0	553	2	US-08-943-087-28	Sequence 28, Appl
c 261	7	1.0	553	2	US-08-943-087-30	Sequence 30, Appl
c 262	7	1.0	553	2	US-08-943-087-32	Sequence 32, Appl
c 263	7	1.0	553	2	US-08-943-087-34	Sequence 34, Appl
c 264	7	1.0	553	2	US-08-943-087-36	Sequence 36, Appl
c 265	7	1.0	553	2	US-08-943-087-38	Sequence 38, Appl
c 266	7	1.0	553	2	US-08-943-087-40	Sequence 40, Appl
c 267	7	1.0	553	2	US-08-943-087-42	Sequence 42, Appl
c 268	7	1.0	553	2	US-08-943-087-44	Sequence 44, Appl
c 269	7	1.0	553	2	US-08-943-087-46	Sequence 46, Appl
c 270	7	1.0	553	2	US-08-943-087-48	Sequence 48, Appl
c 271	7	1.0	553	3	US-09-083-351-2	Sequence 2, Appli
c 272	7	1.0	553	4	US-09-083-352-2	Sequence 2, Appli
273	7	1.0	560	3	US-09-176-657-7	Sequence 7, Appli
274	7	1.0	563	4	US-09-718-841-2	Sequence 2, Appli
275	7	1.0	563	4	US-09-718-810-2	Sequence 2, Appli
276	7	1.0	564	4	US-09-134-001C-3930	Sequence 3930, Ap
c 277	7	1.0	567	3	US-08-646-273-19	Sequence 19, Appl
c 278	7	1.0	567	4	US-09-134-001C-5646	Sequence 5646, Ap
c 279	7	1.0	572	1	US-08-160-861-3	Sequence 3, Appli
c 280	7	1.0	573	3	US-09-295-186-11	Sequence 11, Appl
c 281	7	1.0	582	1	US-08-261-086-2	Sequence 2, Appli
c 282	7	1.0	617	2	US-08-370-156-6	Sequence 6, Appli
c 283	7	1.0	617	3	US-08-814-095-6	Sequence 6, Appli
284	7	1.0	635	1	US-07-832-855-2	Sequence 2, Appli
285	7	1.0	635	4	US-08-176-320-2	Sequence 2, Appli
c 286	7	1.0	652	2	US-08-751-305-2	Sequence 2, Appli
c 287	7	1.0	655	1	US-08-148-910-12	Sequence 12, Appl
c 288	7	1.0	655	1	US-08-448-937A-12	Sequence 12, Appl
c 289	7	1.0	686	2	US-08-993-228-12	Sequence 12, Appl
c 290	7	1.0	695	1	US-08-164-839-8	Sequence 8, Appli
c 291	7	1.0	695	1	US-08-583-799-8	Sequence 8, Appli
c 292	7	1.0	696	1	US-08-164-839-10	Sequence 10, Appl
c 293	7	1.0	696	1	US-08-583-799-10	Sequence 10, Appl
c 294	7	1.0	703	3	US-08-646-273-25	Sequence 25, Appl
c 295	7	1.0	703	4	US-09-134-001C-3015	Sequence 3015, Ap
c 296	7	1.0	708	3	US-08-646-273-23	Sequence 23, Appl
c 297	7	1.0	722	4	US-09-105-058C-23	Sequence 23, Appl
298	7	1.0	723	4	US-09-594-669-14	Sequence 14, Appl
299	7	1.0	725	4	US-09-594-669-16	Sequence 16, Appl
300	7	1.0	730	4	US-09-291-170A-3	Sequence 3, Appli
301	7	1.0	730	4	US-09-724-884-3	Sequence 3, Appli
c 302	7	1.0	753	3	US-08-646-273-36	Sequence 36, Appl
c 303	7	1.0	754	3	US-08-646-273-30	Sequence 30, Appl
c 304	7	1.0	757	4	US-09-177-650-89	Sequence 89, Appl

c 305	7	1.0	758	1	US-08-289-112-2	Sequence 2, Appli
c 306	7	1.0	763	2	US-08-677-862-2	Sequence 2, Appli
c 307	7	1.0	763	2	US-09-252-571-2	Sequence 2, Appli
c 308	7	1.0	763	3	US-09-434-065-2	Sequence 2, Appli
c 309	7	1.0	763	4	US-08-789-275-4	Sequence 4, Appli
c 310	7	1.0	763	4	US-08-789-275-5	Sequence 5, Appli
311	7	1.0	812	4	US-09-166-350-12	Sequence 12, Appl
c 312	7	1.0	814	1	US-08-233-788A-42	Sequence 42, Appl
313	7	1.0	822	2	US-08-939-002A-16	Sequence 16, Appl
c 314	7	1.0	837	4	US-09-390-234-14	Sequence 14, Appl
c 315	7	1.0	863	4	US-09-619-353-14	Sequence 14, Appl
c 316	7	1.0	871	4	US-09-105-058C-20	Sequence 20, Appl
c 317	7	1.0	872	4	US-09-177-650-2	Sequence 2, Appli
c 318	7	1.0	911	4	US-08-461-562B-2	Sequence 2, Appli
c 319	7	1.0	930	4	US-09-177-650-96	Sequence 96, Appl
320	7	1.0	934	1	US-08-446-486-7	Sequence 7, Appli
321	7	1.0	934	1	US-08-463-308-7	Sequence 7, Appli
322	7	1.0	955	1	US-08-006-676B-1	Sequence 1, Appli
323	7	1.0	955	1	US-08-282-845-2	Sequence 2, Appli
324	7	1.0	955	2	US-08-428-414A-3	Sequence 3, Appli
325	7	1.0	955	5	PCT-US94-00324-1	Sequence 1, Appli
c 326	7	1.0	982	4	US-09-556-877-176	Sequence 176, App
c 327	7	1.0	982	4	US-09-620-412C-176	Sequence 176, App
c 328	7	1.0	1006	4	US-09-556-877-190	Sequence 190, App
c 329	7	1.0	1006	4	US-09-620-412C-190	Sequence 190, App
c 330	7	1.0	1093	3	US-08-545-860D-55	Sequence 55, Appl
c 331	7	1.0	1093	5	PCT-US94-04496-55	Sequence 55, Appl
c 332	7	1.0	1164	4	US-09-457-708-2	Sequence 2, Appli
333	7	1.0	1231	4	US-09-071-035-420	Sequence 420, App
334	7	1.0	1232	4	US-09-592-054-2	Sequence 2, Appli
335	7	1.0	1234	4	US-09-592-054-8	Sequence 8, Appli
336	7	1.0	1251	5	PCT-US95-02251-3	Sequence 3, Appli
337	7	1.0	1252	1	US-08-199-780-3	Sequence 3, Appli
338	7	1.0	1252	2	US-08-316-650-3	Sequence 3, Appli
339	7	1.0	1253	3	US-08-479-722B-4	Sequence 4, Appli
340	7	1.0	1265	4	US-09-071-035-418	Sequence 418, App
341	7	1.0	1279	4	US-09-724-517-2	Sequence 2, Appli
342	7	1.0	1279	4	US-09-641-807A-2	Sequence 2, Appli
343	7	1.0	1279	4	US-09-723-096-2	Sequence 2, Appli
c 344	7	1.0	1306	4	US-08-989-299-7	Sequence 7, Appli
c 345	7	1.0	1311	2	US-08-540-406-4	Sequence 4, Appli
c 346	7	1.0	1311	3	US-08-656-055-4	Sequence 4, Appli
c 347	7	1.0	1311	4	US-08-954-668-4	Sequence 4, Appli
c 348	7	1.0	1311	4	US-08-918-658-4	Sequence 4, Appli
c 349	7	1.0	1311	5	PCT-US95-13233-4	Sequence 4, Appli
c 350	7	1.0	1312	4	US-09-041-886-19	Sequence 19, Appl
c 351	7	1.0	1375	4	US-09-210-361-4	Sequence 4, Appli
c 352	7	1.0	1375	4	US-09-740-274-4	Sequence 4, Appli
353	7	1.0	1388	4	US-09-572-191-2	Sequence 2, Appli
354	7	1.0	1388	4	US-09-723-262-2	Sequence 2, Appli
355	7	1.0	1388	4	US-09-723-219-2	Sequence 2, Appli
c 356	7	1.0	1620	1	US-08-542-363-2	Sequence 2, Appli
c 357	7	1.0	1620	4	US-09-100-089-2	Sequence 2, Appli
c 358	7	1.0	1620	4	US-09-670-827-2	Sequence 2, Appli
359	7	1.0	1646	4	US-09-535-008-67	Sequence 67, Appl
360	7	1.0	1647	4	US-09-535-008-2	Sequence 2, Appli
361	7	1.0	1649	4	US-09-535-008-75	Sequence 75, Appl
362	7	1.0	1650	4	US-09-535-008-71	Sequence 71, Appl
c 363	7	1.0	1663	2	US-08-793-126-1	Sequence 1, Appli
c 364	7	1.0	1663	4	US-09-132-271-1	Sequence 1, Appli
c 365	7	1.0	1663	4	US-09-142-334-22	Sequence 22, Appl
366	7	1.0	1678	4	US-09-535-008-69	Sequence 69, Appl
367	7	1.0	1679	4	US-09-535-008-65	Sequence 65, Appl
368	7	1.0	1681	4	US-09-535-008-77	Sequence 77, Appl
369	7	1.0	1682	4	US-09-535-008-73	Sequence 73, Appl
370	7	1.0	1704	4	US-08-485-355B-40	Sequence 40, Appl
371	7	1.0	1872	6	5386025-6	Patent No. 5386025
372	7	1.0	1873	1	US-08-435-675B-4	Sequence 4, Appli
373	7	1.0	1873	1	US-08-336-257A-7	Sequence 7, Appli
374	7	1.0	2318	4	US-09-091-219-24	Sequence 24, Appl
c 375	7	1.0	2368	1	US-08-198-446B-15	Sequence 15, Appl
c 376	7	1.0	2368	2	US-08-870-693-15	Sequence 15, Appl
c 377	7	1.0	3177	2	US-08-477-451-4	Sequence 4, Appli

c	378	6	0.9	8	1	US-08-479-233-10	Sequence 10, Appl
c	379	6	0.9	8	5	PCT-US93-00643-10	Sequence 10, Appl
c	380	6	0.9	9	2	US-08-448-418-88	Sequence 88, Appl
c	381	6	0.9	9	2	US-08-801-319-2	Sequence 2, Appli
	382	6	0.8	9	2	US-08-455-073A-11	Sequence 11, Appl
c	383	6	0.9	10	3	US-08-390-874C-2	Sequence 2, Appli
	384	6	0.8	10	3	US-08-899-595-8	Sequence 8, Appli
	385	6	0.8	10	4	US-09-199-242E-11	Sequence 11, Appl
c	386	6	0.9	10	4	US-09-265-772-2	Sequence 2, Appli
c	387	6	0.9	11	1	US-08-323-474-5	Sequence 5, Appli
c	388	6	0.9	11	1	US-08-086-335C-13	Sequence 13, Appl
c	389	6	0.9	11	1	US-08-641-704-1	Sequence 1, Appli
c	390	6	0.9	11	1	US-08-236-918A-17	Sequence 17, Appl
	391	6	0.8	11	1	US-08-469-582-6	Sequence 6, Appli
	392	6	0.8	11	1	US-08-469-582-7	Sequence 7, Appli
	393	6	0.8	11	2	US-08-473-025-14	Sequence 14, Appl
	394	6	0.8	11	3	US-08-851-843A-186	Sequence 186, App
	395	6	0.8	11	4	US-08-974-549A-305	Sequence 305, App
	396	6	0.8	11	4	US-08-854-050-186	Sequence 186, App
	397	6	0.8	11	4	US-09-430-323-186	Sequence 186, App
c	398	6	0.9	11	4	US-09-150-864A-17	Sequence 17, Appl
c	399	6	0.9	11	5	PCT-US95-08533-1	Sequence 1, Appli
	400	6	0.8	12	1	US-08-469-582-9	Sequence 9, Appli
	401	6	0.8	12	2	US-08-950-866-8	Sequence 8, Appli
	402	6	0.8	12	2	US-08-473-025-13	Sequence 13, Appl
	403	6	0.8	12	2	US-08-224-591-7	Sequence 7, Appli
	404	6	0.8	12	2	US-08-392-338A-1	Sequence 1, Appli
	405	6	0.8	12	2	US-08-926-789-7	Sequence 7, Appli
	406	6	0.8	12	2	US-08-818-253-41	Sequence 41, Appl
	407	6	0.8	12	3	US-09-166-750-1	Sequence 1, Appli
	408	6	0.8	12	3	US-09-166-093-1	Sequence 1, Appli
c	409	6	0.9	12	3	US-08-599-226-31	Sequence 31, Appl
	410	6	0.8	12	3	US-09-172-019-1	Sequence 1, Appli
	411	6	0.8	12	3	US-09-166-094-1	Sequence 1, Appli
	412	6	0.8	12	4	US-08-818-252-41	Sequence 41, Appl
c	413	6	0.9	12	4	US-09-125-098-31	Sequence 31, Appl
	414	6	0.8	12	4	US-09-362-805-3	Sequence 3, Appli
	415	6	0.8	12	4	US-09-173-190-3	Sequence 3, Appli
	416	6	0.8	12	4	US-09-069-821-6	Sequence 6, Appli
	417	6	0.8	12	4	US-09-316-919-57	Sequence 57, Appl
	418	6	0.8	12	5	PCT-US93-11138-7	Sequence 7, Appli
	419	6	0.8	13	1	US-07-763-512-5	Sequence 5, Appli
	420	6	0.8	13	3	US-08-630-916A-7	Sequence 7, Appli
c	421	6	0.9	13	4	US-08-602-999A-78	Sequence 78, Appl
c	422	6	0.9	13	4	US-08-278-865-78	Sequence 78, Appl
	423	6	0.8	13	4	US-08-525-539A-66	Sequence 66, Appl
c	424	6	0.9	13	4	US-09-500-124-78	Sequence 78, Appl
	425	6	0.8	13	4	US-08-469-260A-599	Sequence 599, App
	426	6	0.8	14	1	US-07-994-469A-20	Sequence 20, Appl
	427	6	0.8	14	1	US-08-469-582-3	Sequence 3, Appli
c	428	6	0.9	14	1	US-08-484-969-2	Sequence 2, Appli
c	429	6	0.9	14	1	US-08-472-627-2	Sequence 2, Appli
c	430	6	0.9	14	1	US-08-388-463-2	Sequence 2, Appli
c	431	6	0.9	14	2	US-08-448-418-89	Sequence 89, Appl
	432	6	0.8	14	2	US-08-373-190-3	Sequence 3, Appli
	433	6	0.8	14	2	US-08-373-190-5	Sequence 5, Appli
	434	6	0.8	14	2	US-08-473-025-8	Sequence 8, Appli
	435	6	0.8	14	2	US-08-438-190A-3	Sequence 3, Appli
	436	6	0.8	14	2	US-08-438-190A-5	Sequence 5, Appli
	437	6	0.8	14	2	US-08-818-253-46	Sequence 46, Appl
	438	6	0.8	14	3	US-08-350-215-3	Sequence 3, Appli
	439	6	0.8	14	3	US-08-350-215-5	Sequence 5, Appli
c	440	6	0.9	14	3	US-08-837-226-8	Sequence 8, Appli
	441	6	0.8	14	3	US-09-287-145A-3	Sequence 3, Appli
	442	6	0.8	14	3	US-09-287-145A-5	Sequence 5, Appli
	443	6	0.8	14	4	US-09-046-992-6	Sequence 6, Appli
	444	6	0.8	14	4	US-08-818-252-46	Sequence 46, Appl
	445	6	0.8	14	4	US-08-928-213B-167	Sequence 167, App
	446	6	0.8	14	4	US-09-362-805-8	Sequence 8, Appli
	447	6	0.8	14	4	US-09-173-190-8	Sequence 8, Appli
	448	6	0.8	14	4	US-09-556-111-3	Sequence 3, Appli
	449	6	0.8	14	4	US-09-556-111-5	Sequence 5, Appli
c	450	6	0.9	14	4	US-09-400-653A-33	Sequence 33, Appl
c	451	6	0.9	14	4	US-09-248-061B-17	Sequence 17, Appl
	452	6	0.8	14	4	US-09-316-919-62	Sequence 62, Appl
	453	6	0.8	14	5	PCT-US93-12169-3	Sequence 3, Appli
	454	6	0.8	15	1	US-07-763-512-2	Sequence 2, Appli
	455	6	0.8	15	1	US-08-137-800-5	Sequence 5, Appli
c	456	6	0.9	15	1	US-08-230-047-29	Sequence 29, Appl
	457	6	0.8	15	1	US-08-477-383-5	Sequence 5, Appli
	458	6	0.8	15	1	US-08-487-174-5	Sequence 5, Appli
	459	6	0.8	15	1	US-08-480-750-5	Sequence 5, Appli
	460	6	0.8	15	2	US-08-373-190-4	Sequence 4, Appli
c	461	6	0.9	15	2	US-08-964-725-18	Sequence 18, Appl
	462	6	0.8	15	2	US-08-438-190A-4	Sequence 4, Appli
	463	6	0.8	15	3	US-08-350-215-4	Sequence 4, Appli
c	464	6	0.9	15	3	US-08-654-623-13	Sequence 13, Appl
	465	6	0.8	15	3	US-08-256-747C-26	Sequence 26, Appl
	466	6	0.8	15	3	US-08-630-172-38	Sequence 38, Appl
	467	6	0.8	15	3	US-09-287-145A-4	Sequence 4, Appli
	468	6	0.8	15	4	US-08-834-130A-26	Sequence 26, Appl
	469	6	0.8	15	4	US-08-602-999A-356	Sequence 356, App
	470	6	0.8	15	4	US-08-602-999A-423	Sequence 423, App
	471	6	0.8	15	4	US-09-375-419-38	Sequence 38, Appl
	472	6	0.8	15	4	US-09-556-111-4	Sequence 4, Appli
	473	6	0.8	15	4	US-09-500-124-356	Sequence 356, App
	474	6	0.8	15	4	US-09-500-124-423	Sequence 423, App
	475	6	0.8	15	5	PCT-US93-06751-112	Sequence 112, App
	476	6	0.8	15	5	PCT-US94-01499-4	Sequence 4, Appli
	477	6	0.8	15	5	PCT-US96-07962-2	Sequence 2, Appli
	478	6	0.8	16	1	US-07-994-469A-21	Sequence 21, Appl
c	479	6	0.9	16	2	US-08-637-759B-465	Sequence 465, App
	480	6	0.8	16	2	US-08-859-201-18	Sequence 18, Appl
	481	6	0.8	16	2	US-08-859-201-20	Sequence 20, Appl
	482	6	0.8	16	2	US-08-597-803-2	Sequence 2, Appli
c	483	6	0.9	16	2	US-08-672-345C-43	Sequence 43, Appl
c	484	6	0.9	16	3	US-08-871-355A-465	Sequence 465, App
c	485	6	0.9	16	3	US-08-788-231A-19	Sequence 19, Appl
c	486	6	0.9	16	4	US-09-214-095D-43	Sequence 43, Appl
	487	6	0.8	16	4	US-09-514-302-5	Sequence 5, Appli
c	488	6	0.9	16	4	US-09-201-945-465	Sequence 465, App
c	489	6	0.9	16	6	5196511-27	Patent No. 5196511
	490	6	0.8	17	1	US-07-763-512-3	Sequence 3, Appli
c	491	6	0.9	17	1	US-08-353-400-30	Sequence 30, Appl
c	492	6	0.9	17	1	US-08-467-420A-10	Sequence 10, Appl
c	493	6	0.9	17	1	US-08-470-110A-10	Sequence 10, Appl
c	494	6	0.9	17	1	US-08-667-769A-10	Sequence 10, Appl
c	495	6	0.9	17	2	US-08-940-371-10	Sequence 10, Appl
	496	6	0.8	17	3	US-08-837-226-6	Sequence 6, Appli
c	497	6	0.9	17	3	US-09-015-605-3	Sequence 3, Appli
c	498	6	0.9	17	3	US-08-637-647-10	Sequence 10, Appl
c	499	6	0.9	17	4	US-08-974-549A-61	Sequence 61, Appl
c	500	6	0.9	17	4	US-07-987-264-4	Sequence 4, Appli
c	501	6	0.9	17	4	US-09-265-653-19	Sequence 19, Appl
c	502	6	0.9	17	4	US-09-265-653-22	Sequence 22, Appl
c	503	6	0.9	17	5	PCT-US93-08435-22	Sequence 22, Appl
c	504	6	0.9	17	5	PCT-US93-08435-35	Sequence 35, Appl
c	505	6	0.9	17	5	PCT-US95-17082A-10	Sequence 10, Appl
	506	6	0.8	18	1	US-07-763-512-6	Sequence 6, Appli
	507	6	0.8	18	1	US-07-725-331-33	Sequence 33, Appl
	508	6	0.8	18	1	US-07-725-331-42	Sequence 42, Appl
c	509	6	0.9	18	1	US-08-207-169A-7	Sequence 7, Appli
	510	6	0.8	18	2	US-08-894-339-7	Sequence 7, Appli
	511	6	0.8	18	2	US-08-466-860-20	Sequence 20, Appl
	512	6	0.8	18	3	US-08-472-040A-20	Sequence 20, Appl
	513	6	0.8	18	4	US-09-306-044-7	Sequence 7, Appli
	514	6	0.8	18	4	US-08-276-776-20	Sequence 20, Appl
	515	6	0.8	18	4	US-08-471-209-20	Sequence 20, Appl
	516	6	0.8	18	5	PCT-US91-05047-33	Sequence 33, Appl
	517	6	0.8	18	5	PCT-US91-05047-42	Sequence 42, Appl
c	518	6	0.9	19	1	US-08-457-274A-3	Sequence 3, Appli
c	519	6	0.9	19	2	US-08-652-450A-9	Sequence 9, Appli
c	520	6	0.9	19	2	US-08-448-418-90	Sequence 90, Appl
c	521	6	0.9	19	4	US-08-630-915A-169	Sequence 169, App
c	522	6	0.9	19	4	US-09-225-322B-3	Sequence 3, Appli
c	523	6	0.9	19	5	PCT-US95-05758-3	Sequence 3, Appli

	524	6	0.8	20	1	US-07-955-041-2	Sequence 2, Appli
c	525	6	0.9	20	1	US-08-103-742-32	Sequence 32, Appl
	526	6	0.8	20	1	US-07-956-848A-44	Sequence 44, Appl
	527	6	0.8	20	1	US-08-227-455-2	Sequence 2, Appli
c	528	6	0.9	20	1	US-07-678-974D-35	Sequence 35, Appl
	529	6	0.8	20	1	US-08-472-482-2	Sequence 2, Appli
	530	6	0.8	20	1	US-08-487-069-2	Sequence 2, Appli
	531	6	0.8	20	1	US-08-471-956-44	Sequence 44, Appl
c	532	6	0.9	20	2	US-08-945-168-40	Sequence 40, Appl
c	533	6	0.9	20	3	US-08-654-623-15	Sequence 15, Appl
	534	6	0.8	20	3	US-08-256-747C-12	Sequence 12, Appl
	535	6	0.8	20	4	US-08-834-130A-12	Sequence 12, Appl
c	536	6	0.9	20	4	US-09-196-293-2	Sequence 2, Appli
c	537	6	0.9	20	4	US-08-505-250-22	Sequence 22, Appl
c	538	6	0.9	20	4	US-08-918-288-51	Sequence 51, Appl
c	539	6	0.9	20	4	US-08-918-288-56	Sequence 56, Appl
c	540	6	0.9	20	4	US-09-282-357-51	Sequence 51, Appl
c	541	6	0.9	20	4	US-09-282-357-56	Sequence 56, Appl
c	542	6	0.9	20	4	US-08-209-603E-2	Sequence 2, Appli
c	543	6	0.9	20	4	US-08-505-250-22	Sequence 22, Appl
c	544	6	0.9	21	1	US-07-715-397A-2	Sequence 2, Appli
c	545	6	0.9	21	1	US-07-715-397A-3	Sequence 3, Appli
	546	6	0.8	21	1	US-07-593-657-1	Sequence 1, Appli
c	547	6	0.9	21	1	US-08-060-833-2	Sequence 2, Appli
c	548	6	0.9	21	1	US-08-060-833-3	Sequence 3, Appli
c	549	6	0.9	21	1	US-08-218-025A-124	Sequence 124, App
c	550	6	0.9	21	1	US-08-418-893D-12	Sequence 12, Appl
c	551	6	0.9	21	1	US-08-419-824-2	Sequence 2, Appli
c	552	6	0.9	21	1	US-08-419-824-3	Sequence 3, Appli
c	553	6	0.9	21	2	US-08-845-926-34	Sequence 34, Appl
c	554	6	0.9	21	2	US-08-484-905-14	Sequence 14, Appl
c	555	6	0.9	21	3	US-08-481-985B-14	Sequence 14, Appl
c	556	6	0.9	21	3	US-08-848-580-2	Sequence 2, Appli
c	557	6	0.9	21	3	US-08-848-580-3	Sequence 3, Appli
c	558	6	0.9	21	3	US-08-848-580-6	Sequence 6, Appli
c	559	6	0.9	21	3	US-08-848-580-7	Sequence 7, Appli
c	560	6	0.9	21	3	US-08-848-580-13	Sequence 13, Appl
c	561	6	0.9	21	4	US-08-370-476-14	Sequence 14, Appl
	562	6	0.8	21	4	US-08-602-999A-427	Sequence 427, App
	563	6	0.8	21	4	US-09-500-124-427	Sequence 427, App
c	564	6	0.9	21	4	US-09-351-296-34	Sequence 34, Appl
c	565	6	0.9	21	5	PCT-US92-04537-2	Sequence 2, Appli
c	566	6	0.9	21	5	PCT-US92-04537-3	Sequence 3, Appli
c	567	6	0.9	21	5	PCT-US92-04537-8	Sequence 8, Appli
c	568	6	0.9	21	5	PCT-US92-04537-9	Sequence 9, Appli
	569	6	0.8	21	6	5164481-4	Patent No. 5164481
c	570	6	0.9	22	1	US-08-197-496A-6	Sequence 6, Appli
c	571	6	0.9	22	1	US-08-275-370-6	Sequence 6, Appli
c	572	6	0.9	22	1	US-08-367-968-6	Sequence 6, Appli
c	573	6	0.9	22	2	US-08-665-484-6.	Sequence 6, Appli
c	574	6	0.9	22	2	US-08-792-553-16	Sequence 16, Appl
c	575	6	0.9	22	4	US-09-430-564-9	Sequence 9, Appli
c	576	6	0.9	22	4	US-09-430-564-15	Sequence 15, Appl
	577	6	0.8	23	1	US-08-118-270-93	Sequence 93, Appl
	578	6	0.8	23	1	US-08-268-251-42	Sequence 42, Appl
	579	6	0.8	23	1	US-08-268-251-52	Sequence 52, Appl
c	580	6	0.9	23	1	US-08-393-985-24	Sequence 24, Appl
c	581	6	0.9	23	2	US-08-652-450A-8	Sequence 8, Appli
c	582	6	0.9	23	2	US-08-652-450A-10	Sequence 10, Appl
c	583	6	0.9	23	2	US-08-652-450A-14	Sequence 14, Appl
c	584	6	0.9	23	2	US-08-652-450A-15	Sequence 15, Appl
c	585	6	0.9	23	2	US-08-652-450A-17	Sequence 17, Appl
c	586	6	0.9	23	2	US-08-652-450A-18	Sequence 18, Appl
c	587	6	0.9	23	2	US-08-612-999-2	Sequence 2, Appli
c	588	6	0.9	23	3	US-08-701-382-2	Sequence 2, Appli
	589	6	0.8	23	3	US-08-256-747C-60	Sequence 60, Appl
	590	6	0.8	23	4	US-08-834-130A-60	Sequence 60, Appl
	591	6	0.8	23	4	US-08-557-006C-30	Sequence 30, Appl
c	592	6	0.9	23	4	US-09-028-937-2	Sequence 2, Appli
	593	6	0.8	23	4	US-09-227-357-467	Sequence 467, App
c	594	6	0.9	23	4	US-08-788-820-2	Sequence 2, Appli
	595	6	0.8	23	5	PCT-US93-01112-42	Sequence 42, Appl
	596	6	0.8	23	5	PCT-US93-01112-52	Sequence 52, Appl
	597	6	0.8	23	5	PCT-US93-08528-93	Sequence 93, Appl
c	598	6	0.9	24	3	US-08-592-500-25	Sequence 25, Appl
c	599	6	0.9	24	3	US-08-195-006-25	Sequence 25, Appl
	600	6	0.8	24	4	US-09-098-901-9	Sequence 9, Appli
c	601	6	0.9	24	5	PCT-US94-07644A-25	Sequence 25, Appl
c	602	6	0.9	25	1	US-08-240-514-56	Sequence 56, Appl
c	603	6	0.9	25	2	US-08-612-302A-56	Sequence 56, Appl
c	604	6	0.9	25	2	US-08-902-516-4	Sequence 4, Appli
c	605	6	0.9	25	2	US-08-902-516-10	Sequence 10, Appl
c	606	6	0.9	25	4	US-08-278-865-45	Sequence 45, Appl
c	607	6	0.9	25	4	US-09-305-984-29	Sequence 29, Appl
c	608	6	0.9	25	4	US-09-073-541A-29	Sequence 29, Appl
c	609	6	0.9	25	5	PCT-US94-04361-36	Sequence 36, Appl
c	610	6	0.9	26	1	US-07-942-245-328	Sequence 328, App
	611	6	0.8	26	1	US-07-942-245-425	Sequence 425, App
c	612	6	0.9	26	1	US-07-942-245-487	Sequence 487, App
c	613	6	0.9	26	1	US-07-942-245-490	Sequence 490, App
c	614	6	0.9	26	3	US-08-630-172-40	Sequence 40, Appl
c	615	6	0.9	26	3	US-08-513-968-53	Sequence 53, Appl
c	616	6	0.9	26	4	US-08-602-999A-45	Sequence 45, Appl
	617	6	0.8	26	4	US-08-905-223-289	Sequence 289, App
c	618	6	0.9	26	4	US-09-375-419-40	Sequence 40, Appl
c	619	6	0.9	26	4	US-09-336-536-59	Sequence 59, Appl
c	620	6	0.9	26	4	US-09-500-124-45	Sequence 45, Appl
	621	6	0.8	27	1	US-07-927-071-6	Sequence 6, Appli
	622	6	0.8	27	1	US-07-927-071-7	Sequence 7, Appli
c	623	6	0.9	27	2	US-08-652-450A-6	Sequence 6, Appli
c	624	6	0.9	27	2	US-08-652-450A-7	Sequence 7, Appli
	625	6	0.8	27	2	US-08-833-546-8	Sequence 8, Appli
c	626	6	0.9	27	3	US-08-654-623-23	Sequence 23, Appl
	627	6	0.8	27	3	US-08-256-747C-46	Sequence 46, Appl
	628	6	0.8	27	4	US-08-834-130A-46	Sequence 46, Appl
	629	6	0.8	27	4	US-09-388-664-8	Sequence 8, Appli
c	630	6	0.9	27	4	US-09-227-357-276	Sequence 276, App
c	631	6	0.9	27	6	5166318-17	Patent No. 5166318
c	632	6	0.9	28	1	US-08-164-151-22	Sequence 22, Appl
	633	6	0.8	28	2	US-08-459-568-58	Sequence 58, Appl
	634	6	0.8	28	2	US-08-399-411-58	Sequence 58, Appl
	635	6	0.8	28	3	US-09-045-632-95	Sequence 95, Appl
	636	6	0.8	28	3	US-08-516-859A-58	Sequence 58, Appl
	637	6	0.8	28	4	US-08-957-001B-25	Sequence 25, Appl
	638	6	0.8	28	4	US-09-496-301-25	Sequence 25, Appl
	639	6	0.8	28	4	US-09-586-472-58	Sequence 58, Appl
	640	6	0.8	28	4	US-09-528-706-58	Sequence 58, Appl
c	641	6	0.9	28	5	PCT-US95-04617-2	Sequence 2, Appli
	642	6	0.8	29	1	US-07-729-099-4	Sequence 4, Appli
	643	6	0.8	29	1	US-08-257-392-4	Sequence 4, Appli
c	644	6	0.9	29	1	US-08-393-985-33	Sequence 33, Appl
	645	6	0.8	29	1	US-07-960-510-4	Sequence 4, Appli
	646	6	0.8	29	1	US-07-960-510-5	Sequence 5, Appli
	647	6	0.8	29	1	US-07-960-510-6	Sequence 6, Appli
	648	6	0.8	29	1	US-07-960-510-7	Sequence 7, Appli
	649	6	0.8	29	3	US-08-770-035-4	Sequence 4, Appli
	650	6	0.8	29	3	US-08-256-747C-28	Sequence 28, Appl
	651	6	0.8	29	4	US-08-834-130A-28	Sequence 28, Appl
c	652	6	0.9	30	2	US-08-723-306-25	Sequence 25, Appl
c	653	6	0.9	30	2	US-08-723-306-30	Sequence 30, Appl
c	654	6	0.9	30	4	US-09-348-578-5	Sequence 5, Appli
	655	6	0.8	30	4	US-07-861-458C-132	Sequence 132, App
	656	6	0.8	30	4	US-09-043-816E-14	Sequence 14, Appl
c	657	6	0.9	30	4	US-09-699-684-5	Sequence 5, Appli
c	658	6	0.9	30	5	PCT-US94-04361-31	Sequence 31, Appl
c	659	6	0.9	30	5	PCT-US96-10041-25	Sequence 25, Appl
c	660	6	0.9	30	5	PCT-US96-10041-30	Sequence 30, Appl
	661	6	0.8	31	3	US-07-401-432-52	Sequence 52, Appl
c	662	6	0.9	31	4	US-09-348-578-6	Sequence 6, Appli
c	663	6	0.9	31	4	US-09-348-578-14	Sequence 14, Appl
	664	6	0.8	31	4	US-09-023-731-8	Sequence 8, Appli
	665	6	0.8	31	4	US-09-227-357-491	Sequence 491, App
c	666	6	0.9	31	4	US-09-699-684-6	Sequence 6, Appli
c	667	6	0.9	31	4	US-09-699-684-14	Sequence 14, Appl
	668	6	0.8	32	1	US-07-952-735A-9	Sequence 9, Appli
	669	6	0.8	32	1	US-07-952-735A-10	Sequence 10, Appl

	670	6	0.8	32	1	US-07-952-735A-11	Sequence 11, Appl
	671	6	0.8	32	1	US-08-424-866-3	Sequence 3, Appli
	672	6	0.8	32	1	US-08-468-700-38	Sequence 38, Appl
	673	6	0.8	32	2	US-08-468-220-38	Sequence 38, Appl
	674	6	0.8	32	2	US-08-468-698-38	Sequence 38, Appl
	675	6	0.8	32	2	US-08-704-706A-38	Sequence 38, Appl
	676	6	0.8	32	2	US-08-595-868C-44	Sequence 44, Appl
	677	6	0.8	32	2	US-08-595-868C-45	Sequence 45, Appl
	678	6	0.8	32	2	US-08-595-868C-46	Sequence 46, Appl
c	679	6	0.9	32	3	US-08-938-548B-7	Sequence 7, Appli
	680	6	0.8	32	3	US-07-401-432-53	Sequence 53, Appl
c	681	6	0.9	32	4	US-09-348-578-7	Sequence 7, Appli
c	682	6	0.9	32	4	US-09-348-578-15	Sequence 15, Appl
c	683	6	0.9	32	4	US-09-348-578-23	Sequence 23, Appl
	684	6	0.8	32	4	US-08-828-533-9	Sequence 9, Appli
	685	6	0.8	32	4	US-08-985-659-39	Sequence 39, Appl
	686	6	0.8	32	4	US-09-139-819A-44	Sequence 44, Appl
	687	6	0.8	32	4	US-09-139-819A-45	Sequence 45, Appl
	688	6	0.8	32	4	US-09-139-819A-46	Sequence 46, Appl
	689	6	0.8	32	4	US-08-194-664A-38	Sequence 38, Appl
c	690	6	0.9	32	4	US-08-939-093A-7	Sequence 7, Appli
	691	6	0.8	32	4	US-09-227-357-207	Sequence 207, App
	692	6	0.8	32	4	US-09-750-913-44	Sequence 44, Appl
	693	6	0.8	32	4	US-09-750-913-45	Sequence 45, Appl
	694	6	0.8	32	4	US-09-750-913-46	Sequence 46, Appl
c	695	6	0.9	32	4	US-09-699-684-7	Sequence 7, Appli
c	696	6	0.9	32	4	US-09-699-684-15	Sequence 15, Appl
c	697	6	0.9	32	4	US-09-699-684-23	Sequence 23, Appl
	698	6	0.8	32	5	PCT-US94-01553A-38	Sequence 38, Appl
	699	6	0.8	32	5	PCT-US95-10426-38	Sequence 38, Appl
	700	6	0.8	32	5	PCT-US96-05372-3	Sequence 3, Appli
	701	6	0.8	32	6	5183802-6	Patent No. 5183802
	702	6	0.8	32	6	5183802-7	Patent No. 5183802
	703	6	0.8	32	6	5183802-8	Patent No. 5183802
	704	6	0.8	33	1	US-07-776-272-5	Sequence 5, Appli
	705	6	0.8	33	1	US-07-776-272-6	Sequence 6, Appli
c	706	6	0.9	33	4	US-09-348-578-8	Sequence 8, Appli
c	707	6	0.9	33	4	US-09-348-578-16	Sequence 16, Appl
c	708	6	0.9	33	4	US-09-348-578-24	Sequence 24, Appl
c	709	6	0.9	33	4	US-09-699-684-8	Sequence 8, Appli
c	710	6	0.9	33	4	US-09-699-684-16	Sequence 16, Appl
c	711	6	0.9	33	4	US-09-699-684-24	Sequence 24, Appl
c	712	6	0.9	34	1	US-08-085-122-10	Sequence 10, Appl
	713	6	0.8	34	1	US-08-441-534A-13	Sequence 13, Appl
	714	6	0.8	34	1	US-08-629-752-13	Sequence 13, Appl
c	715	6	0.9	34	1	US-08-623-195-1	Sequence 1, Appli
	716	6	0.8	34	2	US-08-802-991-13	Sequence 13, Appl
c	717	6	0.9	34	2	US-08-902-516-14	Sequence 14, Appl
c	718	6	0.9	34	2	US-08-319-052-20	Sequence 20, Appl
c	719	6	0.9	34	3	US-08-817-335-3	Sequence 3, Appli
c	720	6	0.9	34	3	US-08-906-769-184	Sequence 184, App
c	721	6	0.9	34	3	US-08-906-616-184	Sequence 184, App
c	722	6	0.9	34	4	US-08-639-075A-184	Sequence 184, App
c	723	6	0.9	34	4	US-09-348-578-9	Sequence 9, Appli
c	724	6	0.9	34	4	US-09-348-578-17	Sequence 17, Appl
c	725	6	0.9	34	4	US-09-348-578-25	Sequence 25, Appl
c	726	6	0.9	34	4	US-09-004-731-69	Sequence 69, Appl
c	727	6	0.9	34	4	US-09-012-431-184	Sequence 184, App
c	728	6	0.9	34	4	US-08-749-699-69	Sequence 69, Appl
c	729	6	0.9	34	4	US-09-012-692-184	Sequence 184, App
c	730	6	0.9	34	4	US-08-906-613-184	Sequence 184, App
c	731	6	0.9	34	4	US-08-442-108B-20	Sequence 20, Appl
c	732	6	0.9	34	4	US-09-397-680A-4	Sequence 4, Appli
c	733	6	0.9	34	4	US-09-301-978C-19	Sequence 19, Appl
c	734	6	0.9	34	4	US-09-004-729-69	Sequence 69, Appl
c	735	6	0.9	34	4	US-09-699-684-9	Sequence 9, Appli
c	736	6	0.9	34	4	US-09-699-684-17	Sequence 17, Appl
c	737	6	0.9	34	4	US-09-699-684-25	Sequence 25, Appl
c	738	6	0.9	35	2	US-08-652-450A-5	Sequence 5, Appli
c	739	6	0.9	35	2	US-08-284-391B-47	Sequence 47, Appl
c	740	6	0.9	35	4	US-09-348-578-18	Sequence 18, Appl
c	741	6	0.9	35	4	US-09-348-578-26	Sequence 26, Appl
c	742	6	0.9	35	4	US-09-218-950-47	Sequence 47, Appl
	743	6	0.8	35	4	US-09-315-304B-1466	Sequence 1466, Ap
c	744	6	0.9	35	4	US-09-699-684-18	Sequence 18, Appl
c	745	6	0.9	35	4	US-09-699-684-26	Sequence 26, Appl
	746	6	0.8	35	6	5171673-7	Patent No. 5171673
	747	6	0.8	36	3	US-08-256-747C-25	Sequence 25, Appl
c	748	6	0.9	36	4	US-09-348-578-27	Sequence 27, Appl
	749	6	0.8	36	4	US-08-834-130A-25	Sequence 25, Appl
c	750	6	0.9	36	4	US-09-699-684-27	Sequence 27, Appl
	751	6	0.8	37	3	US-08-545-196B-64	Sequence 64, Appl
c	752	6	0.9	37	4	US-09-227-357-209	Sequence 209, App
c	753	6	0.9	38	4	US-08-525-539A-2	Sequence 2, Appli
c	754	6	0.9	38	4	US-09-336-536-49	Sequence 49, Appl
	755	6	0.8	39	3	US-08-554-385-22	Sequence 22, Appl
	756	6	0.8	39	3	US-08-256-747C-64	Sequence 64, Appl
	757	6	0.8	39	4	US-08-834-130A-64	Sequence 64, Appl
	758	6	0.8	41	1	US-08-865-773-10	Sequence 10, Appl
	759	6	0.8	41	2	US-08-117-952-794	Sequence 794, App
c	760	6	0.9	41	2	US-08-245-511-33	Sequence 33, Appl
c	761	6	0.9	41	2	US-08-600-993A-33	Sequence 33, Appl
	762	6	0.8	41	4	US-08-981-189B-6	Sequence 6, Appli
	763	6	0.8	41	4	US-09-400-716-8	Sequence 8, Appli
	764	6	0.8	41	4	US-09-424-127-10	Sequence 10, Appl
	765	6	0.8	43	3	US-08-256-747C-27	Sequence 27, Appl
	766	6	0.8	43	4	US-08-974-549A-202	Sequence 202, App
	767	6	0.8	43	4	US-08-834-130A-27	Sequence 27, Appl
	768	6	0.8	43	4	US-08-750-624-14	Sequence 14, Appl
c	769	6	0.9	43	4	US-07-757-022B-12	Sequence 12, Appl
c	770	6	0.9	44	1	US-08-262-037-133	Sequence 133, App
	771	6	0.8	44	4	US-09-227-357-220	Sequence 220, App
	772	6	0.8	45	3	US-08-256-747C-65	Sequence 65, Appl
	773	6	0.8	45	4	US-08-834-130A-65	Sequence 65, Appl
c	774	6	0.9	46	4	US-08-483-533-30	Sequence 30, Appl
c	775	6	0.9	46	4	US-09-283-471A-30	Sequence 30, Appl
c	776	6	0.9	46	4	US-07-757-022B-18	Sequence 18, Appl
c	777	6	0.9	48	4	US-09-004-406C-15	Sequence 15, Appl
	778	6	0.8	48	4	US-09-288-143-102	Sequence 102, App
	779	6	0.8	50	1	US-08-321-071A-31	Sequence 31, Appl
	780	6	0.8	50	1	US-08-464-339A-16	Sequence 16, Appl
c	781	6	0.9	50	3	US-08-630-916A-122	Sequence 122, App
c	782	6	0.9	50	4	US-08-905-223-488	Sequence 488, App
	783	6	0.8	51	2	US-08-760-075A-27	Sequence 27, Appl
c	784	6	0.9	51	4	US-08-444-818-2	Sequence 2, Appli
	785	6	0.8	51	4	US-09-338-546-27	Sequence 27, Appl
c	786	6	0.9	51	4	US-09-605-785-567	Sequence 567, App
	787	6	0.8	51	4	US-09-659-084-27	Sequence 27, Appl
c	788	6	0.9	51	4	US-07-757-022B-100	Sequence 100, App
c	789	6	0.9	51	4	US-07-757-022B-112	Sequence 112, App
c	790	6	0.9	52	2	US-08-993-228-4	Sequence 4, Appli
c	791	6	0.9	52	4	US-09-153-599A-4	Sequence 4, Appli
c	792	6	0.9	53	1	US-08-222-616-6	Sequence 6, Appli
c	793	6	0.9	53	4	US-08-446-648-6	Sequence 6, Appli
c	794	6	0.9	53	5	PCT-US95-04228-6	Sequence 6, Appli
c	795	6	0.9	54	1	US-08-497-134A-16	Sequence 16, Appl
c	796	6	0.9	54	1	US-08-464-339A-15	Sequence 15, Appl
	797	6	0.8	54	3	US-08-851-843A-19	Sequence 19, Appl
	798	6	0.8	54	4	US-08-854-050-19	Sequence 19, Appl
	799	6	0.8	54	4	US-09-430-323-19	Sequence 19, Appl
c	800	6	0.9	55	4	US-09-361-707-91	Sequence 91, Appl
c	801	6	0.9	55	4	US-09-361-707-93	Sequence 93, Appl
c	802	6	0.9	55	4	US-09-361-707-95	Sequence 95, Appl
c	803	6	0.9	55	4	US-09-361-707-96	Sequence 96, Appl
c	804	6	0.9	55	4	US-09-361-707-97	Sequence 97, Appl
c	805	6	0.9	55	4	US-09-361-707-99	Sequence 99, Appl
c	806	6	0.9	55	4	US-09-361-707-100	Sequence 100, App
c	807	6	0.9	57	2	US-08-318-837-50	Sequence 50, Appl
c	808	6	0.9	59	2	US-08-414-657D-52	Sequence 52, Appl
c	809	6	0.9	59	2	US-08-414-657D-53	Sequence 53, Appl
	810	6	0.8	59	4	US-09-488-799-29	Sequence 29, Appl
c	811	6	0.9	60	4	US-09-227-357-227	Sequence 227, App
c	812	6	0.9	61	4	US-08-936-165A-408	Sequence 408, App
	813	6	0.8	62	4	US-09-227-357-231	Sequence 231, App
	814	6	0.8	64	1	US-07-689-693B-13	Sequence 13, Appl
	815	6	0.8	64	1	US-08-137-800-36	Sequence 36, Appl

	816	6	0.8	64	1	US-08-137-800-45	Sequence 45, Appl
	817	6	0.8	64	1	US-08-477-383-36	Sequence 36, Appl
	818	6	0.8	64	1	US-08-477-383-45	Sequence 45, Appl
	819	6	0.8	64	1	US-08-487-174-36	Sequence 36, Appl
	820	6	0.8	64	1	US-08-487-174-45	Sequence 45, Appl
	821	6	0.8	64	1	US-08-480-750-36	Sequence 36, Appl
	822	6	0.8	64	1	US-08-480-750-45	Sequence 45, Appl
c	823	6	0.9	65	4	US-09-134-001C-3146	Sequence 3146, Ap
c	824	6	0.9	67	4	US-09-003-198A-16	Sequence 16, Appl
c	825	6	0.9	67	4	US-09-134-001C-5251	Sequence 5251, Ap
	826	6	0.8	67	4	US-09-149-476-430	Sequence 430, App
c	827	6	0.9	68	2	US-08-385-590A-20	Sequence 20, Appl
c	828	6	0.9	68	3	US-09-021-520-20	Sequence 20, Appl
c	829	6	0.9	68	4	US-09-134-001C-5173	Sequence 5173, Ap
c	830	6	0.9	69	1	US-08-358-160-107	Sequence 107, App
c	831	6	0.9	69	2	US-08-385-590A-18	Sequence 18, Appl
c	832	6	0.9	69	3	US-09-021-520-18	Sequence 18, Appl
c	833	6	0.9	69	4	US-09-134-001C-3787	Sequence 3787, Ap
	834	6	0.8	70	4	US-08-905-223-290	Sequence 290, App
	835	6	0.8	70	4	US-09-367-953B-70	Sequence 70, Appl
	836	6	0.8	70	4	US-09-367-953B-74	Sequence 74, Appl
	837	6	0.8	70	4	US-09-367-953B-78	Sequence 78, Appl
	838	6	0.8	70	4	US-09-367-953B-96	Sequence 96, Appl
c	839	6	0.9	70	4	US-09-006-428A-15	Sequence 15, Appl
c	840	6	0.9	71	2	US-08-812-003-9	Sequence 9, Appli
c	841	6	0.9	71	4	US-09-134-001C-2906	Sequence 2906, Ap
	842	6	0.8	71	4	US-08-469-260A-207	Sequence 207, App
	843	6	0.8	73	2	US-08-891-837B-2	Sequence 2, Appli
	844	6	0.8	73	4	US-09-134-001C-5316	Sequence 5316, Ap
c	845	6	0.9	74	1	US-08-543-238-2	Sequence 2, Appli
c	846	6	0.9	74	1	US-08-543-238-5	Sequence 5, Appli
c	847	6	0.9	74	1	US-08-420-526-2	Sequence 2, Appli
c	848	6	0.9	74	1	US-08-420-526-5	Sequence 5, Appli
	849	6	0.8	74	1	US-08-594-031-97	Sequence 97, Appl
	850	6	0.8	74	4	US-08-866-545-6	Sequence 6, Appli
	851	6	0.8	75	3	US-08-256-747C-56	Sequence 56, Appl
	852	6	0.8	75	4	US-08-834-130A-56	Sequence 56, Appl
	853	6	0.8	75	4	US-09-134-001C-5507	Sequence 5507, Ap
c	854	6	0.9	76	5	PCT-US95-13658-6	Sequence 6, Appli
c	855	6	0.9	78	4	US-09-130-339-4	Sequence 4, Appli
c	856	6	0.9	79	2	US-08-633-879C-14	Sequence 14, Appl
c	857	6	0.9	79	2	US-09-174-060-1	Sequence 1, Appli
c	858	6	0.9	79	3	US-08-338-382-1	Sequence 1, Appli
c	859	6	0.9	79	4	US-09-100-802-4	Sequence 4, Appli
	860	6	0.8	80	1	US-08-011-398B-14	Sequence 14, Appl
	861	6	0.8	80	1	US-08-011-398B-15	Sequence 15, Appl
	862	6	0.8	80	1	US-08-011-398B-16	Sequence 16, Appl
	863	6	0.8	80	1	US-08-464-051-14	Sequence 14, Appl
	864	6	0.8	80	1	US-08-464-051-15	Sequence 15, Appl
	865	6	0.8	80	1	US-08-464-051-16	Sequence 16, Appl
	866	6	0.8	80	2	US-08-462-498-14	Sequence 14, Appl
	867	6	0.8	80	2	US-08-462-498-15	Sequence 15, Appl
	868	6	0.8	80	2	US-08-462-498-16	Sequence 16, Appl
c	869	6	0.9	80	3	US-08-492-459-16	Sequence 16, Appl
	870	6	0.8	80	3	US-08-554-385-13	Sequence 13, Appl
	871	6	0.8	80	3	US-08-554-385-14	Sequence 14, Appl
	872	6	0.8	80	3	US-08-554-385-15	Sequence 15, Appl
c	873	6	0.9	80	3	US-08-423-752-16	Sequence 16, Appl
c	874	6	0.9	80	4	US-08-716-873-30	Sequence 30, Appl
c	875	6	0.9	80	4	US-08-905-223-411	Sequence 411, App
c	876	6	0.9	80	4	US-09-368-431-30	Sequence 30, Appl
c	877	6	0.9	80	4	US-09-414-006-16	Sequence 16, Appl
c	878	6	0.9	80	4	US-09-447-223-16	Sequence 16, Appl
c	879	6	0.9	81	2	US-08-248-839C-146	Sequence 146, App
c	880	6	0.9	81	2	US-08-245-511-14	Sequence 14, Appl
c	881	6	0.9	81	2	US-08-600-993A-14	Sequence 14, Appl
c	882	6	0.9	81	4	US-09-100-802-5	Sequence 5, Appli
c	883	6	0.9	82	1	US-08-497-312-30	Sequence 30, Appl
	884	6	0.8	82	4	US-09-513-783A-136	Sequence 136, App
	885	6	0.8	83	2	US-08-726-306A-76	Sequence 76, Appl
c	886	6	0.9	84	4	US-09-386-493-8	Sequence 8, Appli
c	887	6	0.9	84	4	US-09-386-493-14	Sequence 14, Appl
	888	6	0.8	87	4	US-09-299-495F-31	Sequence 31, Appl
	889	6	0.8	88	2	US-08-690-011A-31	Sequence 31, Appl
	890	6	0.8	88	3	US-09-045-632-7	Sequence 7, Appli
c	891	6	0.9	88	4	US-09-330-970-6	Sequence 6, Appli
c	892	6	0.9	88	4	US-09-263-933-23	Sequence 23, Appl
c	893	6	0.9	89	4	US-09-214-631-13	Sequence 13, Appl
c	894	6	0.9	91	2	US-08-997-080-143	Sequence 143, App
c	895	6	0.9	91	2	US-08-997-362-143	Sequence 143, App
c	896	6	0.9	91	4	US-09-095-855-143	Sequence 143, App
c	897	6	0.9	91	4	US-09-324-542-143	Sequence 143, App
c	898	6	0.9	91	4	US-09-134-001C-5117	Sequence 5117, Ap
c	899	6	0.9	91	4	US-09-205-426-143	Sequence 143, App
	900	6	0.8	92	1	US-08-290-448A-68	Sequence 68, Appl
	901	6	0.8	92	1	US-08-290-448A-68	Sequence 68, Appl
	902	6	0.8	92	1	US-08-175-069A-68	Sequence 68, Appl
	903	6	0.8	92	4	US-08-461-939B-68	Sequence 68, Appl
	904	6	0.8	92	4	US-08-464-000-68	Sequence 68, Appl
c	905	6	0.9	92	6	5166318-12	Patent No. 5166318
c	906	6	0.9	93	2	US-08-964-725-14	Sequence 14, Appl
c	907	6	0.9	94	4	US-08-469-260A-170	Sequence 170, App
	908	6	0.8	94	4	US-08-055-006-23	Sequence 23, Appl
c	909	6	0.9	95	4	US-09-188-930-344	Sequence 344, App
c	910	6	0.9	95	4	US-09-724-864-68	Sequence 68, Appl
	911	6	0.8	95	4	US-09-134-001C-5541	Sequence 5541, Ap
	912	6	0.8	96	3	US-09-101-167-10	Sequence 10, Appl
	913	6	0.8	97	2	US-08-242-097-5	Sequence 5, Appli
	914	6	0.8	97	4	US-09-206-695-5	Sequence 5, Appli
	915	6	0.8	97	4	US-09-309-487-23	Sequence 23, Appl
	916	6	0.8	97	4	US-09-227-357-465	Sequence 465, App
c	917	6	0.9	97	4	US-09-422-936-73	Sequence 73, Appl
c	918	6	0.9	97	4	US-09-422-936-83	Sequence 83, Appl
c	919	6	0.9	98	2	US-08-825-556A-4	Sequence 4, Appli
c	920	6	0.9	98	4	US-09-134-001C-4312	Sequence 4312, Ap
c	921	6	0.9	99	2	US-08-825-556A-3	Sequence 3, Appli
c	922	6	0.9	99	2	US-08-484-905-62	Sequence 62, Appl
c	923	6	0.9	99	3	US-08-481-985B-62	Sequence 62, Appl
c	924	6	0.9	99	4	US-09-188-930-340	Sequence 340, App
c	925	6	0.9	99	4	US-08-370-476-62	Sequence 62, Appl
c	926	6	0.9	100	4	US-08-679-493A-146	Sequence 146, App
	927	6	0.8	102	2	US-08-808-982-8	Sequence 8, Appli
c	928	6	0.9	102	2	US-09-174-060-2	Sequence 2, Appli
c	929	6	0.9	102	3	US-08-338-382-2	Sequence 2, Appli
	930	6	0.8	102	4	US-09-306-902A-8	Sequence 8, Appli
c	931	6	0.9	102	4	US-09-134-001C-4995	Sequence 4995, Ap
	932	6	0.8	103	2	US-08-529-878B-41	Sequence 41, Appl
c	933	6	0.9	103	2	US-08-529-878B-41	Sequence 41, Appl
c	934	6	0.9	103	3	US-08-783-974-1	Sequence 1, Appli
c	935	6	0.9	105	1	US-08-467-420A-16	Sequence 16, Appl
c	936	6	0.9	105	1	US-08-470-110A-16	Sequence 16, Appl
c	937	6	0.9	105	2	US-08-940-371-16	Sequence 16, Appl
c	938	6	0.9	105	3	US-08-488-551B-639	Sequence 639, App
c	939	6	0.9	105	3	US-08-637-647-16	Sequence 16, Appl
c	940	6	0.9	105	4	US-09-001-472-2	Sequence 2, Appli
c	941	6	0.9	105	4	US-09-422-936-81	Sequence 81, Appl
c	942	6	0.9	106	1	US-08-352-324A-5	Sequence 5, Appli
c	943	6	0.9	106	2	US-08-862-607-5	Sequence 5, Appli
c	944	6	0.9	106	2	US-08-468-819-7	Sequence 7, Appli
c	945	6	0.9	106	3	US-08-946-329A-76	Sequence 76, Appl
c	946	6	0.9	106	3	US-09-203-235-5	Sequence 5, Appli
c	947	6	0.9	106	4	US-08-679-493A-148	Sequence 148, App
c	948	6	0.9	106	4	US-09-189-129-4	Sequence 4, Appli
c	949	6	0.9	106	4	US-08-858-207A-326	Sequence 326, App
c	950	6	0.9	106	5	PCT-US95-16144-5	Sequence 5, Appli
c	951	6	0.9	107	1	US-08-352-324A-4	Sequence 4, Appli
c	952	6	0.9	107	1	US-08-352-324A-7	Sequence 7, Appli
c	953	6	0.9	107	2	US-08-862-607-4	Sequence 4, Appli
c	954	6	0.9	107	2	US-08-862-607-7	Sequence 7, Appli
c	955	6	0.9	107	2	US-08-468-819-5	Sequence 5, Appli
c	956	6	0.9	107	2	US-08-468-819-6	Sequence 6, Appli
c	957	6	0.9	107	3	US-09-203-235-4	Sequence 4, Appli
c	958	6	0.9	107	3	US-09-203-235-7	Sequence 7, Appli
c	959	6	0.9	107	5	PCT-US95-16144-4	Sequence 4, Appli
c	960	6	0.9	107	5	PCT-US95-16144-7	Sequence 7, Appli
c	961	6	0.9	108	1	US-08-468-661-3	Sequence 3, Appli

c 962	6	0.9	108	1	US-08-466-272A-3	Sequence 3, Appli
c 963	6	0.9	108	1	US-08-478-857-3	Sequence 3, Appli
c 964	6	0.9	108	1	US-08-488-113B-151	Sequence 151, App
c 965	6	0.9	108	1	US-08-477-484B-151	Sequence 151, App
c 966	6	0.9	108	1	US-08-107-669D-15	Sequence 15, Appl
c 967	6	0.9	108	1	US-08-472-788A-15	Sequence 15, Appl
c 968	6	0.9	108	2	US-08-477-531B-15	Sequence 15, Appl
c 969	6	0.9	108	2	US-08-646-360-151	Sequence 151, App
c 970	6	0.9	108	2	US-08-471-771-3	Sequence 3, Appli
c 971	6	0.9	108	2	US-08-082-842A-15	Sequence 15, Appl
c 972	6	0.9	108	3	US-08-388-353-639	Sequence 639, App
c 973	6	0.9	108	3	US-09-130-783-3	Sequence 3, Appli
974	6	0.8	108	3	US-08-959-212-10	Sequence 10, Appl
c 975	6	0.9	108	4	US-08-839-765-151	Sequence 151, App
c 976	6	0.9	108	4	US-09-136-389-151	Sequence 151, App
977	6	0.8	108	4	US-08-936-165A-443	Sequence 443, App
c 978	6	0.9	108	4	US-09-199-637A-227	Sequence 227, App
c 979	6	0.9	108	4	US-09-610-838-151	Sequence 151, App
980	6	0.8	109	2	US-08-210-762E-22	Sequence 22, Appl
c 981	6	0.9	109	2	US-08-308-494A-23	Sequence 23, Appl
982	6	0.8	109	4	US-09-106-075A-22	Sequence 22, Appl
983	6	0.8	110	4	US-09-134-001C-3110	Sequence 3110, Ap
c 984	6	0.9	110	4	US-09-134-001C-3110	Sequence 3110, Ap
985	6	0.8	110	4	US-09-134-001C-5023	Sequence 5023, Ap
986	6	0.8	110	6	5514582-9	Patent No. 5514582
c 987	6	0.9	111	1	US-08-466-886-19	Sequence 19, Appl
c 988	6	0.9	111	2	US-08-825-556A-2	Sequence 2, Appli
c 989	6	0.9	111	2	US-08-800-198-6	Sequence 6, Appli
c 990	6	0.9	111	3	US-09-296-595-6	Sequence 6, Appli
c 991	6	0.9	111	4	US-08-469-617-19	Sequence 19, Appl
992	6	0.8	111	4	US-09-347-801-8	Sequence 8, Appli
993	6	0.8	111	4	US-09-134-001C-3859	Sequence 3859, Ap
c 994	6	0.9	112	1	US-08-053-171-16	Sequence 16, Appl
c 995	6	0.9	112	1	US-08-331-398A-49	Sequence 49, Appl
c 996	6	0.9	112	1	US-08-477-877B-87	Sequence 87, Appl
c 997	6	0.9	112	1	US-08-477-877B-88	Sequence 88, Appl
998	6	0.8	112	1	US-08-446-919A-5	Sequence 5, Appli
c 999	6	0.9	112	2	US-08-472-281A-87	Sequence 87, Appl
c1000	6	0.9	112	2	US-08-472-281A-88	Sequence 88, Appl

ALIGNMENTS

```
RESULT 1
US-09-058-489-15
; Sequence 15, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 15
;   LENGTH: 662
;   TYPE: PRT
;   ORGANISM: Human
US-09-058-489-15

Alignment Scores:
Pred. No.:              1.7e-12       Length:         662
Score:                  22.00         Matches:        22
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            3.04%         Indels:         0
DB:                     3             Gaps:           0

US-09-714-865-15 (1-2172) x US-09-058-489-15 (1-662)

Qy         976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         218 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 237

Qy        1036 ATTTTG 1041
               ||||||
Db         238 IleLeu 239

RESULT 2
US-09-058-489-16
; Sequence 16, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
;   LENGTH: 662
;   TYPE: PRT
;   ORGANISM: Human
US-09-058-489-16

Alignment Scores:
Pred. No.:              1.7e-12       Length:         662
Score:                  22.00         Matches:        22
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            3.04%         Indels:         0
DB:                     3             Gaps:           0

US-09-714-865-15 (1-2172) x US-09-058-489-16 (1-662)

Qy         976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         218 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 237

Qy        1036 ATTTTG 1041
               ||||||
Db         238 IleLeu 239

RESULT 3
US-09-058-489-18
; Sequence 18, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
;   LENGTH: 660
;   TYPE: PRT
;   ORGANISM: Human
```

```
US-09-058-489-18

Alignment Scores:
Pred. No.:                1.03e-08      Length:         660
Score:                    18.00         Matches:        18
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              2.49%         Indels:         0
DB:                       3             Gaps:           0

US-09-714-865-15 (1-2172) x US-09-058-489-18 (1-660)

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTG 1041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeu 237

RESULT 4
US-09-058-489-91
; Sequence 91, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 91
;   LENGTH: 660
;   TYPE: PRT
;   ORGANISM: Human
US-09-058-489-91

Alignment Scores:
Pred. No.:                1.03e-08      Length:         660
Score:                    18.00         Matches:        18
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              2.49%         Indels:         0
DB:                       3             Gaps:           0

US-09-714-865-15 (1-2172) x US-09-058-489-91 (1-660)

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTG 1041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeu 237

RESULT 5
US-09-183-706-43
; Sequence 43, Application US/09183706
; Patent No. 6245525
; GENERAL INFORMATION:
;  APPLICANT: Martelange, Valrie
;  APPLICANT: De Smet, Charles
;  APPLICANT: Boon-Falleur, Thierry
;  TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE: L0461/7054
;  CURRENT APPLICATION NUMBER: US/09/183,706
;  CURRENT FILING DATE: 1998-10-30
;  EARLIER APPLICATION NUMBER: 09/122,989
;  EARLIER FILING DATE: 1999-07-27
;  NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 43
;   LENGTH: 648
;   TYPE: PRT
;   ORGANISM: H. sapiens
US-09-183-706-43

Alignment Scores:
Pred. No.:                0.0432        Length:         648
Score:                    11.00         Matches:        11
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              1.52%         Indels:         0
DB:                       4             Gaps:           0

US-09-714-865-15 (1-2172) x US-09-183-706-43 (1-648)

Qy       1867 GAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
              |||||||||||||||||||||||||||||||||
Db        570 GluTyrValHisArgIleGlyArgThrGlyArg 580

RESULT 6
US-09-567-995-43
; Sequence 43, Application US/09567995
; Patent No. 6303756
; GENERAL INFORMATION:
;  APPLICANT:  Martelange, Val,rie
;  APPLICANT:  De Smet, Charles
;  APPLICANT:  Boon-Falleur, Thierry
;  TITLE OF INVENTION:  TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE:  L0461/7054
;  CURRENT APPLICATION NUMBER: US/09/567,995
;  CURRENT FILING DATE:  2000-05-10
;  PRIOR APPLICATION NUMBER:  09/183,706
;  PRIOR FILING DATE:  1998-10-30
;  NUMBER OF SEQ ID NOS:  43
; SEQ ID NO 43
;   LENGTH: 648
;   TYPE: PRT
;   ORGANISM: H. sapiens
US-09-567-995-43

Alignment Scores:
Pred. No.:                0.0432        Length:         648
Score:                    11.00         Matches:        11
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              1.52%         Indels:         0
DB:                       4             Gaps:           0

US-09-714-865-15 (1-2172) x US-09-567-995-43 (1-648)

Qy       1867 GAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
              |||||||||||||||||||||||||||||||||
Db        570 GluTyrValHisArgIleGlyArgThrGlyArg 580

RESULT 7
US-08-679-493A-85
; Sequence 85, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
;  APPLICANT: Taylor, Ethan W.
;  TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
;  FILE REFERENCE: 55-95
;  CURRENT APPLICATION NUMBER: US/08/679,493A
;  CURRENT FILING DATE:  1996-07-12
;  PRIOR APPLICATION NUMBER: 60/001203
;  PRIOR FILING DATE: 1995-07-14
;  PRIOR APPLICATION NUMBER: 60/003,112
;  PRIOR FILING DATE: 1995-09-01
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-08-679-493A-85

Alignment Scores:
```

```
Pred. No.:                0.469        Length:        149
Score:                    10.00        Matches:       10
Percent Similarity:       100.00%      Conservative:  0
Best Local Similarity:    100.00%      Mismatches:    0
Query Match:              1.38%        Indels:        0
DB:                       4            Gaps:          0

US-09-714-865-15 (1-2172) x US-08-679-493A-85 (1-149)

Qy       1870 TATGTTCATCGAATTGGGCGTACTGGTCGT 1899
              ||||||||||||||||||||||||||||||
Db         52 TyrValHisArgIleGlyArgThrGlyArg 61

RESULT 8
US-09-393-627B-25
; Sequence 25, Application US/09393627B
; Patent No. 6455314
; GENERAL INFORMATION:
;  APPLICANT: Wickham, Thomas J.
;  APPLICANT:  Kovesdi, Imre
;  APPLICANT:  Roelvink, Petrus W.
;  APPLICANT:  Bruder, Joseph T.
;  TITLE OF INVENTION: Alternatively Targeted Adenovirus
;  FILE REFERENCE: 202345
;  CURRENT APPLICATION NUMBER: US/09/393,627B
;  CURRENT FILING DATE:  1999-09-10
;  PRIOR APPLICATION NUMBER: US 60/099,851
;  PRIOR FILING DATE: 1998-09-11
;  PRIOR APPLICATION NUMBER: US 60/136,529
;  PRIOR FILING DATE: 1999-05-28
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 25
;   LENGTH: 191
;   TYPE: PRT
;   ORGANISM: Human adenovirus serotype 35
US-09-393-627B-25

Alignment Scores:
Pred. No.:                4            Length:        191
Score:                    9.00         Matches:       9
Percent Similarity:       100.00%      Conservative:  0
Best Local Similarity:    100.00%      Mismatches:    0
Query Match:              1.28%        Indels:        0
DB:                       4            Gaps:          0

US-09-714-865-15 (1-2172) x US-09-393-627B-25 (1-191)

Qy        480 AAATCCTCCACGGCAACCTCGGAAACT 454
              |||||||||||||||||||||||||||
Db         86 LysSerSerThrAlaThrSerGluThr 94

RESULT 9
US-08-224-482-4
; Sequence 4, Application US/08224482
; Patent No. 5837692
;  GENERAL INFORMATION:
;    APPLICANT:  Mercola, Dan
;    APPLICANT:  Adamson, Eileen D.
;    TITLE OF INVENTION:  Inhibition of the Mitogenic Activity of
;    TITLE OF INVENTION:  PDGF by Mammalian EGR
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/224,482
;      FILING DATE:  07-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ME 9913
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  543 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-224-482-4

Alignment Scores:
Pred. No.:                3.45         Length:        543
Score:                    9.00         Matches:       9
Percent Similarity:       100.00%      Conservative:  0
Best Local Similarity:    100.00%      Mismatches:    0
Query Match:              1.28%        Indels:        0
DB:                       2            Gaps:          0

US-09-714-865-15 (1-2172) x US-08-224-482-4 (1-543)

Qy        767 TCCTCATCCTCAGGTGGAGGAGGGGGT 741
              |||||||||||||||||||||||||||
Db         64 SerSerSerSerGlyGlyGlyGlyGly 72

RESULT 10
US-09-149-934-4
; Sequence 4, Application US/09149934B
; Patent No. 6139837
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Shah, Purvi
;  TITLE OF INVENTION: ATP-DEPENDENT RNA HELICASE PROTEIN
;  FILE REFERENCE: PF-0338-1 DIV
;  CURRENT APPLICATION NUMBER: US/09/149,934B
;  CURRENT FILING DATE: 1998-09-09
;  EARLIER APPLICATION NUMBER: 08/892,256
;  EARLIER FILING DATE: 1997-07-11
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PERL Program
; SEQ ID NO 4
;   LENGTH: 746
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No. 6139837 1707046
US-09-149-934-4

Alignment Scores:
Pred. No.:                3.3          Length:        746
Score:                    9.00         Matches:       9
Percent Similarity:       100.00%      Conservative:  0
Best Local Similarity:    100.00%      Mismatches:    0
Query Match:              1.24%        Indels:        0
DB:                       4            Gaps:          0

US-09-714-865-15 (1-2172) x US-09-149-934-4 (1-746)

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTC 1134
              |||||||||||||||||||||||||||
Db        240 IleValAlaProThrArgGluLeuVal 248
```

```
RESULT 11
US-09-010-928B-2
; Sequence 2, Application US/09010928B
; Patent No. 5994099
;  GENERAL INFORMATION:
;    APPLICANT:  Lewis, Randolph V
;    APPLICANT:  Hayashi, Cheryl Y
;    TITLE OF INVENTION:  EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
;    TITLE OF INVENTION:  CODING THEREFOR
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  8110 GATEHOUSE RD. SUITE 500E
;      CITY:  FALLS CHURCH
;      STATE:  VIRGINIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/010,928B
;      FILING DATE:  22-JAN-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M
;      REGISTRATION NUMBER:  28977
;      REFERENCE/DOCKET NUMBER:  1447-109P
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  870 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-010-928B-2

Alignment Scores:
Pred. No.:              3.23           Length:         870
Score:                  9.00           Matches:        9
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            1.28%          Indels:         0
DB:                     2              Gaps:           0

US-09-714-865-15 (1-2172) x US-09-010-928B-2 (1-870)

Qy        767 TCCTCATCCTCAGGTGGAGGAGGGGGT 741
              |||||||||||||||||||||||||||
Db         92 SerSerSerSerGlyGlyGlyGlyGly 100

RESULT 12
US-09-208-742-4
; Sequence 4, Application US/09208742
; Patent No. 6174679
; GENERAL INFORMATION:
;  APPLICANT: Kaufmann, Joerg
;  TITLE OF INVENTION: CIF150/hTAFII150 is Necessary for Cell
;  TITLE OF INVENTION: Cycle Progression
;  FILE REFERENCE: 1453.002
;  CURRENT APPLICATION NUMBER: US/09/208,742
;  CURRENT FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 1261
;   TYPE: PRT
;   ORGANISM: human
US-09-208-742-4

Alignment Scores:
Pred. No.:              3.06           Length:         1261
Score:                  9.00           Matches:        9
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            1.24%          Indels:         0
DB:                     4              Gaps:           0

US-09-714-865-15 (1-2172) x US-09-208-742-4 (1-1261)

Qy       1000 ACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||||||||||||||||||||||||||
Db        344 ThrGlySerGlyLysThrAlaAlaPhe 352

RESULT 13
US-09-332-295-2
; Sequence 2, Application US/09332295
; Patent No. 6303372
; GENERAL INFORMATION:
;  APPLICANT: Kaufmann, Joerg
;  TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
;  FILE REFERENCE: 200130.456 / 1513.003
;  CURRENT APPLICATION NUMBER: US/09/332,295
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1261
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-332-295-2

Alignment Scores:
Pred. No.:              3.06           Length:         1261
Score:                  9.00           Matches:        9
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            1.24%          Indels:         0
DB:                     4              Gaps:           0

US-09-714-865-15 (1-2172) x US-09-332-295-2 (1-1261)

Qy       1000 ACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||||||||||||||||||||||||||
Db        344 ThrGlySerGlyLysThrAlaAlaPhe 352

RESULT 14
US-09-709-979-2
; Sequence 2, Application US/09709979
; Patent No. 6423822
; GENERAL INFORMATION:
;  APPLICANT: Kaufmann, Joerg
;  TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
;  FILE REFERENCE: 200130.456 / 1513.003
;  CURRENT APPLICATION NUMBER: US/09/709,979
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: US 09/332,295
;  PRIOR FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1261
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-709-979-2

Alignment Scores:
Pred. No.:              3.06           Length:         1261
Score:                  9.00           Matches:        9
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            1.24%          Indels:         0
DB:                     4              Gaps:           0

US-09-714-865-15 (1-2172) x US-09-709-979-2 (1-1261)
```

```
Qy       1000 ACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||||||||||||||||||||||||||
Db        344 ThrGlySerGlyLysThrAlaAlaPhe 352

RESULT 15
US-08-264-093-24
; Sequence 24, Application US/08264093
; Patent No. 5639863
;  GENERAL INFORMATION:
;    APPLICANT:  Michael D. Dan
;    TITLE OF INVENTION:  HUMAN MONOCLONAL ANTIBODIES SPECIFIC TO
;    TITLE OF INVENTION:  CELL CYCLE-INDEPENDENT GLIOMA SURFACE
;    TITLE OF INVENTION:  ANTIGEN
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ridout & Maybee
;      STREET:  2300 Richmond-Adelaide Centre
;      STREET:  101 Richmond Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 2J7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.4 Mb storage
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  MS-DOS 6.00
;      SOFTWARE:  ASCII Editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/264,093
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:  No. 5639863 applicable
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lake, James R.
;      REGISTRATION NUMBER:  31081
;      REFERENCE/DOCKET NUMBER:  NOVOP/106A/7551
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 868-1482
;      TELEFAX:  (416) 362-0823
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not applicable
;      TOPOLOGY:  linear
US-08-264-093-24

Alignment Scores:
Pred. No.:              49.7          Length:         17
Score:                  8.00          Matches:        8
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            1.14%         Indels:         0
DB:                     1             Gaps:           0

US-09-714-865-15 (1-2172) x US-08-264-093-24 (1-17)

Qy        356 TCTTCGCAGTCATTACTAGACTCT 333
              ||||||||||||||||||||||||
Db          2 SerSerGlnSerLeuLeuAspSer 9


Search completed: June 10, 2003, 17:18:27
Job time : 42.5 secs
```

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:         June 10, 2003, 16:44:46 ; Search time 53.5 Seconds
                                          (without alignments)
                                          10819.452 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  724
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  OLIGO
                Xgapop 60.0 , Xgapext 60.0
                Ygapop 60.0 , Ygapext 60.0
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       908470 seqs, 133250620 residues

Word size:      1

Total number of hits satisfying chosen parameters:      1687582

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111832_26561/app_query.fasta_1.2311
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=oligo.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=1000 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_76_@runat_05062003_111832_26561 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :      A_Geneseq_101002:*
                1:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	724	100.0	724	22	AAE02417	Human vasa protein
2	123	17.0	135	22	AAU23109	Novel human enzyme
3	84	11.6	106	22	ABB95898	Human testicular a
4	84	11.6	106	22	AAU23664	Novel human enzyme
5	84	11.6	106	22	AAM95199	Human reproductive
6	82	11.3	713	22	AAE02419	Rat vasa protein.
7	64	8.8	722	22	AAE02418	Mouse vasa protein
8	30	4.1	30	22	AAE02423	Epitope #1 of huma
9	29	4.0	29	22	AAE02424	Epitope #2 of huma
10	24	3.3	700	22	AAE02420	Frog vasa protein.
11	24	3.3	700	22	AAE02421	Danio reio vasa pr
12	23	3.2	661	22	ABB59954	Drosophila melanog
13	23	3.2	661	22	AAE02422	Fruit fly vasa pro
14	22	3.0	112	21	AAB38176	Gene 6 human secre
15	22	3.0	112	21	AAB38177	Human secreted pro
16	22	3.0	662	19	AAW81502	Dead Box X (DBX) g
17	22	3.0	662	19	AAW81501	Dead Box X (DBX) g
18	22	3.0	662	21	AAY96483	Human DBX1, an RNA
19	22	3.0	662	23	ABB57045	Mouse ischaemic co
20	19	2.6	646	21	AAG17458	Arabidopsis thalia
21	19	2.6	798	22	ABB64631	Drosophila melanog
22	18	2.5	660	19	AAW81503	Dead Box Y (DBY) g
23	17	2.3	612	21	AAG17433	Arabidopsis thalia
24	17	2.3	612	21	AAG20655	Arabidopsis thalia
25	16	2.2	344	21	AAY96482	Human topoisomeras
26	15	2.1	168	23	ABP05062	Human ORFX protein
27	15	2.1	257	21	AAY51577	T. thermophila Hc1
28	15	2.1	343	21	AAB43994	Human cancer assoc
29	15	2.1	500	22	AAE10164	RNA helicase conse
30	15	2.1	574	22	ABB67213	Drosophila melanog
31	15	2.1	577	22	ABB64981	Drosophila melanog
32	15	2.1	578	22	ABB65733	Drosophila melanog
33	15	2.1	578	22	ABB67212	Drosophila melanog
34	15	2.1	614	22	ABB44571	Mouse wound healin
35	15	2.1	614	22	ABB44572	Human wound healin
36	15	2.1	615	22	ABB44573	Mouse wound healin
37	15	2.1	703	22	ABB63551	Drosophila melanog
38	15	2.1	811	22	ABB64859	Drosophila melanog
39	15	2.1	945	22	ABB65231	Drosophila melanog
40	14	1.9	421	21	AAG17434	Arabidopsis thalia
41	14	1.9	421	21	AAG20656	Arabidopsis thalia
42	14	1.9	460	21	AAG17459	Arabidopsis thalia
43	14	1.9	732	22	ABG10453	Novel human diagno
44	13	1.8	149	23	AAU99918	Human 49875 DEAD t
45	13	1.8	457	23	AAU72977	Neisseria meningit
46	12	1.7	136	23	ABP07185	Human ORFX protein
47	12	1.7	222	22	AAU23649	Novel human enzyme
48	12	1.7	399	21	AAG17435	Arabidopsis thalia
49	12	1.7	399	21	AAG20657	Arabidopsis thalia
50	12	1.7	438	21	AAG17460	Arabidopsis thalia
51	12	1.7	619	22	AAG64953	Human ATP-dependen
52	12	1.7	626	22	AAU14731	Novel bone marrow
53	12	1.7	639	22	AAU23092	Novel human enzyme
54	11	1.5	45	22	AAM65383	Human brain expres
55	11	1.5	45	23	ABG47079	Human peptide enco
56	11	1.5	518	22	ABG25120	Novel human diagno
57	11	1.5	589	19	AAW60667	E.coli cold shock
58	11	1.5	613	22	AAU35397	Haemophilus influe
59	11	1.5	646	22	AAU34718	E. coli cellular p
60	11	1.5	646	22	AAU38176	Salmonella typhi c
61	11	1.5	648	20	AAY49636	Human sdp3.8 (HAGE
62	11	1.5	960	22	ABG25032	Novel human diagno
63	11	1.5	1006	22	ABG24871	Novel human diagno
64	11	1.5	1008	22	ABG18393	Novel human diagno
65	10	1.4	68	15	AAR46082	DEAD ATP helicase
66	10	1.4	78	23	ABP35000	Human helicase-lik
67	10	1.4	149	23	AAM52405	RNA helicase fragm
c 68	10	1.4	220	21	AAY75382	Neisseria gonorrhe

69	10	1.4	221	22	AAU23651	Novel human enzyme
70	10	1.4	226	22	AAU23095	Novel human enzyme
71	10	1.4	303	21	AAG46347	Arabidopsis thalia
72	10	1.4	323	22	ABG10452	Novel human diagno
73	10	1.4	327	21	AAG46346	Arabidopsis thalia
74	10	1.4	421	22	AAU36110	Klebsiella pneumon
75	10	1.4	424	22	AAG90605	C glutamicum prote
76	10	1.4	440	22	AAU23549	Novel human enzyme
77	10	1.4	456	21	AAG46345	Arabidopsis thalia
78	10	1.4	460	22	ABG08818	Novel human diagno
79	10	1.4	470	23	ABB48320	Listeria monocytog
80	10	1.4	484	22	AAM25668	Human protein sequ
81	10	1.4	495	23	AAE20093	Lactobacillus rham
82	10	1.4	497	20	AAY41671	Lactobacillus reut
83	10	1.4	520	23	ABB47602	Listeria monocytog
84	10	1.4	524	22	AAU37898	Streptococcus pneu
85	10	1.4	528	23	ABP28778	Streptococcus poly
86	10	1.4	528	23	ABP29908	Streptococcus poly
87	10	1.4	536	22	AAU35181	Enterococcus faeca
88	10	1.4	539	23	ABP28779	Streptococcus poly
89	10	1.4	547	23	ABB53657	Lactococcus lactis
90	10	1.4	567	22	AAU36292	Pseudomonas aerugi
91	10	1.4	568	22	AAU53624	Propionibacterium
92	10	1.4	619	22	ABB69630	Drosophila melanog
93	10	1.4	622	22	AAB92773	Human protein sequ
94	10	1.4	622	22	AAB93391	Human protein sequ
95	10	1.4	624	23	AAE14675	Corynebacterium gl
96	10	1.4	628	21	AAB42879	Human ORFX ORF2643
97	10	1.4	635	20	AAY73847	Human prostate tum
98	10	1.4	648	22	ABG22292	Novel human diagno
99	10	1.4	732	22	AAG91024	C glutamicum prote
100	10	1.4	1038	22	ABG24868	Novel human diagno
101	10	1.4	1224	22	ABB62191	Drosophila melanog
c 102	9	1.3	40	21	AAB38560	Human secreted pro
103	9	1.2	59	22	AAU23800	Novel human enzyme
104	9	1.2	59	22	AAM83264	Human immune/haema
c 105	9	1.3	108	21	AAG34179	Zea mays protein f
c 106	9	1.3	126	21	AAY56558	Flagelliform fibro
c 107	9	1.3	191	21	AAY82522	Human adenovirus t
108	9	1.2	257	22	AAB92816	Human protein sequ
c 109	9	1.3	262	22	ABB66342	Drosophila melanog
c 110	9	1.3	287	22	ABB66343	Drosophila melanog
111	9	1.2	312	23	ABB04728	Human PP3241 prote
c 112	9	1.3	325	22	AAB48515	Fiber protein of A
c 113	9	1.3	337	21	AAY68657	Amino acid sequenc
c 114	9	1.3	338	21	AAY68656	Amino acid sequenc
c 115	9	1.3	338	22	AAB31741	Amino acid sequenc
c 116	9	1.3	338	22	AAB31742	Amino acid sequenc
c 117	9	1.3	338	22	AAB67184	Mastadenovirus 34
c 118	9	1.3	338	22	AAB67185	Mastadenovirus 35
c 119	9	1.3	342	23	AAU79745	Amino acid sequenc
120	9	1.2	419	22	AAU23554	Novel human enzyme
121	9	1.2	441	23	ABB48633	Listeria monocytog
c 122	9	1.3	448	22	ABB58601	Drosophila melanog
123	9	1.2	486	22	ABB66093	Drosophila melanog
c 124	9	1.3	505	22	AAB61923	Human early growth
c 125	9	1.3	543	13	AAR24022	Human promyelo-leu
c 126	9	1.3	543	20	AAW76985	Human EGR-1 protei
c 127	9	1.3	543	21	AAY51116	Human EGR-1 protei
c 128	9	1.3	543	22	AAU09066	Human ZIF268 prote
c 129	9	1.3	543	22	AAB76861	Human lung tumour
c 130	9	1.3	543	23	AAU85516	Clone #18996 of lu
c 131	9	1.3	548	22	ABB59192	Drosophila melanog
132	9	1.2	556	22	AAB73639	Arabidopsis thalia
133	9	1.2	574	23	ABB97289	Novel human protei
c 134	9	1.3	650	22	ABB65739	Drosophila melanog
c 135	9	1.3	663	18	AAW29150	Dual-specific muri
136	9	1.2	703	22	ABG19359	Novel human diagno
137	9	1.2	709	22	AAB76859	Human lung tumour
138	9	1.2	709	23	AAU85514	Clone #18929 of lu
139	9	1.2	746	22	AAB37465	ATP-dependent RNA
140	9	1.2	765	22	AAB73637	Human ATP-dependen
141	9	1.2	765	22	AAB94373	Human protein sequ
142	9	1.2	768	21	AAB42448	Human ORFX ORF2212
143	9	1.2	782	22	ABB58923	Drosophila melanog
144	9	1.2	791	22	ABB62354	Drosophila melanog
c 145	9	1.3	924	21	AAY52991	Drosophila melanog
c 146	9	1.3	958	22	ABB62764	Drosophila melanog
147	9	1.2	1275	20	AAY29084	Human CIF130 prote
148	9	1.2	1275	21	AAY44262	Human cell cycle r
c 149	9	1.3	3502	22	ABB58382	Drosophila melanog
c 150	8	1.1	16	13	AAR24704	Sequence of an ant
c 151	8	1.1	16	22	AAB86292	Murine derived ant
c 152	8	1.1	16	22	AAB73652	Murine monoclonal
c 153	8	1.1	16	23	AAO14984	Mouse mAb 1D9 anti
c 154	8	1.1	20	21	AAY70796	Murine anti-p53 PA
c 155	8	1.1	25	17	AAR88856	Peptide targetted
c 156	8	1.1	39	22	ABB44289	Peptide #11795 enc
c 157	8	1.1	39	22	ABB27151	Protein #9150 enco
c 158	8	1.1	39	22	AAM65326	Human brain expres
c 159	8	1.1	39	22	AAM78023	Human bone marrow
c 160	8	1.1	39	22	AAM21912	Peptide #8346 enco
c 161	8	1.1	39	22	AAM38240	Peptide #12277 enc
c 162	8	1.1	39	23	ABG47038	Human peptide enco
c 163	8	1.1	40	21	AAG56574	Arabidopsis thalia
c 164	8	1.1	53	23	ABP09771	Human ORFX protein
c 165	8	1.1	58	22	AAU22441	Human cardiovascul
166	8	1.1	61	23	ABP34492	Human helicase-lik
c 167	8	1.1	63	21	AAG50020	Arabidopsis thalia
c 168	8	1.1	64	22	AAM96330	Human reproductive
c 169	8	1.1	65	22	AAE06085	Human gene 45 enco
c 170	8	1.1	65	23	ABG33907	Human secreted pro
c 171	8	1.1	66	21	AAY87108	Human secreted pro
c 172	8	1.1	69	23	ABP09715	Human ORFX protein
c 173	8	1.1	72	20	AAY06406	Human secreted pro
c 174	8	1.1	72	22	AAU39017	Human secreted pro
c 175	8	1.1	72	23	ABB55726	Human polypeptide
c 176	8	1.1	73	23	ABP05652	Human ORFX protein
c 177	8	1.1	75	22	AAU30402	Novel human secret
c 178	8	1.1	75	22	AAO10920	Human polypeptide
179	8	1.1	78	16	AAR70161	Streptococcus pneu
180	8	1.1	93	23	AAU91110	Human secreted pro
c 181	8	1.1	95	21	AAY87195	Human secreted pro
c 182	8	1.1	95	22	AAE06172	Human gene 45 enco
c 183	8	1.1	95	23	ABG33996	Human secreted pro
c 184	8	1.1	99	21	AAG09859	Arabidopsis thalia
c 185	8	1.1	99	21	AAG25224	Arabidopsis thalia
c 186	8	1.1	99	21	AAG50019	Arabidopsis thalia
c 187	8	1.1	100	22	AAE06958	Murine 1D9 antibod
c 188	8	1.1	100	22	AAE06959	Mouse germline kap
c 189	8	1.1	110	21	AAY70788	Murine anti-p53 mo
c 190	8	1.1	110	22	AAM77590	Human bone marrow
c 191	8	1.1	110	23	AAO14974	Humanised murine 1
c 192	8	1.1	110	23	ABG46625	Human peptide enco
c 193	8	1.1	111	23	AAO14975	Humanised murine 1
c 194	8	1.1	111	23	AAO14977	Humanised murine 1
c 195	8	1.1	112	13	AAR24710	Sequence of a chim
c 196	8	1.1	112	19	AAW48248	A77 anti-Fc alpha
c 197	8	1.1	112	20	AAY52765	Anti-tissue factor
c 198	8	1.1	112	20	AAY52766	Anti-tissue factor
c 199	8	1.1	112	20	AAW89177	Anti-p53 monoclona
c 200	8	1.1	112	22	AAU09918	Murine mAb 1D9 lig
c 201	8	1.1	112	22	AAU09921	Humanised 1D9 ligh
c 202	8	1.1	112	22	AAU09922	Humanised 1D9 ligh
c 203	8	1.1	112	22	AAU09923	Humanised 1D9 ligh
c 204	8	1.1	112	22	AAU09924	Humanised 1D9 ligh
c 205	8	1.1	112	22	AAU09925	Humanised 1D9 ligh
c 206	8	1.1	112	22	AAE08554	Murine A77 anti-Fc
c 207	8	1.1	112	22	AAE06946	Murine 1D9 antibod
c 208	8	1.1	112	22	AAE06949	Humanised murine 1
c 209	8	1.1	112	22	AAE06950	Humanised murine 1
c 210	8	1.1	112	22	AAE06951	Humanised murine 1
c 211	8	1.1	112	22	AAE06952	Humanised murine 1
c 212	8	1.1	112	22	AAE06991	Human kappa light
c 213	8	1.1	112	22	AAE07036	Humanised murine 1
c 214	8	1.1	112	22	AAB74621	A77 anti-Fc-alpha-

c	215	8	1.1	112	23	AAO14971	Mouse light chain
c	216	8	1.1	112	23	AAO14973	Humanised murine l
c	217	8	1.1	112	23	AAO14976	Humanised murine l
c	218	8	1.1	113	17	AAR88508	VLkappa for antibo
c	219	8	1.1	113	22	AAE07005	Human kappa light
c	220	8	1.1	113	22	AAE03751	Murine PSCA Ab lig
c	221	8	1.1	114	22	AAE07035	Humanised murine a
	222	8	1.1	116	22	ABG30374	Novel human diagno
c	223	8	1.1	118	22	ABG16698	Novel human diagno
	224	8	1.1	119	23	AAU91087	Human secreted pro
c	225	8	1.1	123	22	ABG04683	Novel human diagno
	226	8	1.1	126	23	AAU91137	Human secreted pro
c	227	8	1.1	131	12	AAR12239	Mouse MAb 4D12 L c
c	228	8	1.1	131	22	AAO10924	Human polypeptide
c	229	8	1.1	132	12	AAR12361	Light (kappa) chai
c	230	8	1.1	132	13	AAR24712	Sequence encoded b
c	231	8	1.1	133	21	AAB43731	Human cancer assoc
	232	8	1.1	138	22	ABG25114	Novel human diagno
c	233	8	1.1	142	22	AAE07032	Murine antibody 1D
	234	8	1.1	143	19	AAW98594	H. pylori GHPO 115
	235	8	1.1	143	22	AAO01800	Human polypeptide
c	236	8	1.1	145	22	ABB68802	Drosophila melanog
c	237	8	1.1	152	21	AAG09858	Arabidopsis thalia
	238	8	1.1	158	22	ABG30370	Novel human diagno
c	239	8	1.1	160	21	AAG09857	Arabidopsis thalia
c	240	8	1.1	160	22	AAB35101	Antibody variable
c	241	8	1.1	161	22	AAM24370	Human EST encoded
c	242	8	1.1	166	21	AAG17261	Arabidopsis thalia
c	243	8	1.1	166	21	AAG25223	Arabidopsis thalia
c	244	8	1.1	166	21	AAG42947	Arabidopsis thalia
	245	8	1.1	166	23	ABP39434	Staphylococcus epi
	246	8	1.1	168	22	ABG07782	Novel human diagno
	247	8	1.1	168	22	ABG30371	Novel human diagno
	248	8	1.1	187	20	AAY60035	Human endometrium
c	249	8	1.1	189	21	AAG25222	Arabidopsis thalia
c	250	8	1.1	191	21	AAY82520	Human adenovirus t
c	251	8	1.1	203	21	AAG21890	Arabidopsis thalia
c	252	8	1.1	204	21	AAG19965	Arabidopsis thalia
c	253	8	1.1	205	21	AAG21889	Arabidopsis thalia
	254	8	1.1	205	22	ABG07253	Novel human diagno
c	255	8	1.1	206	21	AAG19964	Arabidopsis thalia
c	256	8	1.1	207	21	AAG19963	Arabidopsis thalia
c	257	8	1.1	210	21	AAG21888	Arabidopsis thalia
	258	8	1.1	211	21	AAB54371	Human pancreatic c
	259	8	1.1	213	21	AAG04742	Arabidopsis thalia
	260	8	1.1	214	22	AAU23539	Novel human enzyme
	261	8	1.1	216	21	AAG04741	Arabidopsis thalia
c	262	8	1.1	218	22	AAE03756	Chimeric antibody
	263	8	1.1	220	21	AAB53303	Human colon cancer
	264	8	1.1	220	21	AAG31903	Arabidopsis thalia
	265	8	1.1	220	22	AAG73751	Human colon cancer
c	266	8	1.1	223	21	AAG14416	Arabidopsis thalia
	267	8	1.1	224	20	AAY60124	Human endometrium
c	268	8	1.1	233	22	AAU27773	Human full-length
c	269	8	1.1	233	22	AAB95172	Human protein sequ
c	270	8	1.1	233	22	AAB36593	Human FLEXHT-15 pr
c	271	8	1.1	236	21	AAB56832	Human prostate can
c	272	8	1.1	243	19	AAW60769	Single chain antib
	273	8	1.1	248	22	ABG08222	Novel human diagno
	274	8	1.1	259	21	AAG04740	Arabidopsis thalia
	275	8	1.1	261	23	ABP43023	Human ovarian anti
c	276	8	1.1	271	18	AAW12692	Natural killer cel
c	277	8	1.1	271	21	AAY43399	Human natural kill
c	278	8	1.1	271	22	AAM50249	Human natural kill
c	279	8	1.1	271	22	AAB82478	Natural killer cel
c	280	8	1.1	271	22	AAB68038	Amino acid sequenc
	281	8	1.1	275	21	AAG35471	Arabidopsis thalia
	282	8	1.1	283	21	AAG31902	Arabidopsis thalia
c	283	8	1.1	296	20	AAY14121	Raly protein seque
c	284	8	1.1	302	21	AAG43670	Arabidopsis thalia
	285	8	1.1	304	21	AAG31901	Arabidopsis thalia
c	286	8	1.1	307	22	ABB58806	Drosophila melanog
c	287	8	1.1	312	21	AAG17260	Arabidopsis thalia
c	288	8	1.1	312	21	AAG42946	Arabidopsis thalia
	289	8	1.1	314	22	ABG22291	Novel human diagno
c	290	8	1.1	320	21	AAG17259	Arabidopsis thalia
c	291	8	1.1	320	21	AAG42945	Arabidopsis thalia
c	292	8	1.1	320	22	AAE02556	A. thaliana transc
c	293	8	1.1	320	23	AAU93119	Arabidopsis transc
c	294	8	1.1	325	22	AAB48516	Fiber protein of A
	295	8	1.1	328	22	AAU37895	Streptococcus pneu
	296	8	1.1	328	22	AAU37963	Streptococcus pneu
	297	8	1.1	328	22	ABG07252	Novel human diagno
c	298	8	1.1	336	21	AAG43669	Arabidopsis thalia
c	299	8	1.1	346	23	AAU11818	Cancer and neuroge
c	300	8	1.1	349	21	AAG14415	Arabidopsis thalia
	301	8	1.1	352	22	ABG07250	Novel human diagno
c	302	8	1.1	353	20	AAY06273	Anti Fc alpha rece
c	303	8	1.1	353	21	AAY68670	Amino acid sequenc
c	304	8	1.1	353	22	AAB31755	Amino acid sequenc
c	305	8	1.1	353	22	AAB67198	Mastadenovirus 51
	306	8	1.1	356	21	AAG35470	Arabidopsis thalia
c	307	8	1.1	357	23	AAU79746	Amino acid sequenc
	308	8	1.1	359	21	AAG43944	Zea mays protein f
	309	8	1.1	361	21	AAG43943	Zea mays protein f
	310	8	1.1	363	21	AAG43942	Zea mays protein f
c	311	8	1.1	367	23	ABB10101	Human homeobox pro
c	312	8	1.1	368	23	AAU11273	Murine beta1,3-N-a
c	313	8	1.1	369	21	AAG41236	Arabidopsis thalia
c	314	8	1.1	371	23	AAU11819	Cancer and neuroge
	315	8	1.1	374	22	AAM52309	Murine zyxine frag
	316	8	1.1	380	22	AAM52311	Human zyxine fragm
	317	8	1.1	387	21	AAB56899	Human prostate can
c	318	8	1.1	391	22	AAM25280	Human protein sequ
	319	8	1.1	401	22	ABG08143	Novel human diagno
	320	8	1.1	407	22	AAB93496	Human protein sequ
	321	8	1.1	413	22	AAB92764	Human protein sequ
	322	8	1.1	413	22	AAB94038	Human protein sequ
c	323	8	1.1	418	21	AAG41235	Arabidopsis thalia
c	324	8	1.1	439	21	AAG41234	Arabidopsis thalia
c	325	8	1.1	440	22	AAG92622	C glutamicum prote
c	326	8	1.1	440	22	AAB79004	C. glutamicum SRT
c	327	8	1.1	442	22	AAB94926	Human protein sequ
	328	8	1.1	443	22	AAM40069	Human polypeptide
	329	8	1.1	447	23	ABP41581	Human ovarian anti
	330	8	1.1	457	21	AAG54360	Arabidopsis thalia
	331	8	1.1	459	21	AAB43582	Human cancer assoc
	332	8	1.1	461	21	AAB42956	Human ORFX ORF2720
	333	8	1.1	461	21	AAG54359	Arabidopsis thalia
c	334	8	1.1	462	22	AAU30404	Novel human secret
	335	8	1.1	468	21	AAG53119	Arabidopsis thalia
	336	8	1.1	471	23	ABG65636	Human breast speci
	337	8	1.1	473	22	ABB64634	Drosophila melanog
	338	8	1.1	478	22	AAB93520	Human protein sequ
	339	8	1.1	478	23	ABB57344	Mouse ischaemic co
	340	8	1.1	479	21	AAG47287	Arabidopsis thalia
	341	8	1.1	479	21	AAY58616	Protein regulating
	342	8	1.1	479	22	AAM39539	Human polypeptide
	343	8	1.1	479	22	AAB93928	Human protein sequ
	344	8	1.1	479	22	AAB37464	Murine ATP-depende
	345	8	1.1	483	23	AAM47315	Human ATP dependen
	346	8	1.1	485	21	AAG53118	Arabidopsis thalia
	347	8	1.1	487	21	AAG48622	Arabidopsis thalia
	348	8	1.1	487	22	AAU36804	Staphylococcus aur
	349	8	1.1	487	22	AAU37398	Staphylococcus aur
	350	8	1.1	489	21	AAG53117	Arabidopsis thalia
	351	8	1.1	491	22	AAB93174	Human protein sequ
	352	8	1.1	492	22	AAU35715	Helicobacter pylor
	353	8	1.1	493	22	ABG07251	Novel human diagno
	354	8	1.1	494	21	AAG54358	Arabidopsis thalia
	355	8	1.1	499	21	AAG35469	Arabidopsis thalia
	356	8	1.1	502	21	AAG13871	Arabidopsis thalia
	357	8	1.1	502	21	AAG46881	Arabidopsis thalia
	358	8	1.1	510	22	AAU33760	Staphylococcus aur
	359	8	1.1	511	22	AAB67687	Amino acid sequenc
	360	8	1.1	514	22	AAM41325	Human polypeptide

	361	8	1.1	521	22	ABB65234	Drosophila melanog
c	362	8	1.1	524	22	ABB64765	Drosophila melanog
c	363	8	1.1	526	22	AAM78355	Human protein SEQ
	364	8	1.1	526	22	AAU00864	S. aureus ATP-depe
	365	8	1.1	527	20	AAY05468	S. aureus dbpA pro
	366	8	1.1	528	23	ABP39417	Staphylococcus epi
	367	8	1.1	529	22	AAU33292	Novel human secret
c	368	8	1.1	535	18	AAW28491	Human p53 protein
c	369	8	1.1	535	18	AAW28492	Human p53 protein
c	370	8	1.1	536	22	AAM78356	Human protein SEQ
	371	8	1.1	547	21	AAY56849	Human RNA binding
	372	8	1.1	547	22	AAM39379	Human polypeptide
	373	8	1.1	547	22	AAB95154	Human protein sequ
	374	8	1.1	547	22	AAB67685	Amino acid sequenc
	375	8	1.1	548	20	AAW88549	Secreted protein e
	376	8	1.1	548	22	ABB50316	Human secreted pro
	377	8	1.1	550	21	AAG48621	Arabidopsis thalia
c	378	8	1.1	552	22	ABB71493	Drosophila melanog
c	379	8	1.1	552	22	ABB71495	Drosophila melanog
	380	8	1.1	564	22	AAM52303	Murine zyxine. Mu
	381	8	1.1	564	23	ABB57260	Mouse ischaemic co
	382	8	1.1	566	22	AAM41165	Human polypeptide
c	383	8	1.1	569	22	AAM79339	Human protein SEQ
c	384	8	1.1	569	22	AAM79340	Human protein SEQ
	385	8	1.1	571	21	AAG48620	Arabidopsis thalia
	386	8	1.1	572	22	AAG68191	Zyxin protein SEQ
	387	8	1.1	572	22	AAM52305	Human zyxine. Hom
	388	8	1.1	574	22	AAM41855	Human polypeptide
c	389	8	1.1	588	21	AAG39789	Arabidopsis thalia
	390	8	1.1	590	22	ABG22143	Novel human diagno
	391	8	1.1	593	21	AAG13870	Arabidopsis thalia
	392	8	1.1	593	21	AAG46880	Arabidopsis thalia
	393	8	1.1	594	21	AAG47286	Arabidopsis thalia
	394	8	1.1	594	22	ABB61659	Drosophila melanog
	395	8	1.1	599	23	AAE15278	Human RNA metaboli
	396	8	1.1	600	23	AAU99903	Human 49875 DEAD t
c	397	8	1.1	603	22	ABB68633	Drosophila melanog
	398	8	1.1	610	22	ABG19757	Novel human diagno
	399	8	1.1	612	21	AAG13869	Arabidopsis thalia
	400	8	1.1	612	21	AAG46879	Arabidopsis thalia
	401	8	1.1	613	22	ABB64540	Drosophila melanog
	402	8	1.1	615	22	ABG07784	Novel human diagno
	403	8	1.1	615	22	ABG09334	Novel human diagno
c	404	8	1.1	619	21	AAG45697	Arabidopsis thalia
c	405	8	1.1	627	20	AAY05851	Banana ripening fr
	406	8	1.1	633	20	AAY44017	Saccharomyces cere
	407	8	1.1	657	21	AAG17814	Arabidopsis thalia
	408	8	1.1	657	21	AAG49207	Arabidopsis thalia
	409	8	1.1	663	22	ABB62293	Drosophila melanog
	410	8	1.1	663	22	ABG30377	Novel human diagno
	411	8	1.1	667	21	AAG17813	Arabidopsis thalia
	412	8	1.1	667	21	AAG49206	Arabidopsis thalia
	413	8	1.1	670	21	AAY70229	Human RNA-associat
	414	8	1.1	670	22	AAM78950	Human protein SEQ
	415	8	1.1	670	22	AAB92989	Human protein sequ
	416	8	1.1	671	21	AAG17812	Arabidopsis thalia
	417	8	1.1	671	21	AAG49205	Arabidopsis thalia
c	418	8	1.1	684	21	AAB10556	Human aspartate pr
	419	8	1.1	684	22	AAM79934	Human protein SEQ
	420	8	1.1	691	21	AAG47285	Arabidopsis thalia
c	421	8	1.1	707	23	AAO14204	Human transporter
c	422	8	1.1	717	23	AAU85408	Human protein NOV6
	423	8	1.1	727	22	ABB71717	Drosophila melanog
c	424	8	1.1	750	21	AAG45696	Arabidopsis thalia
	425	8	1.1	752	20	AAW97799	Streptococcus pneu
	426	8	1.1	752	20	AAW74407	C. trachomatis gid
	427	8	1.1	752	22	AAU23535	Novel human enzyme
	428	8	1.1	752	22	AAB84212	Amino acid sequenc
	429	8	1.1	752	22	AAB84213	Amino acid sequenc
	430	8	1.1	752	23	AAG79339	Glycogen phosphory
	431	8	1.1	764	23	AAU02756	Barley DEAD box pr
c	432	8	1.1	768	21	AAG45695	Arabidopsis thalia
	433	8	1.1	783	22	AAM39072	Human polypeptide
c	434	8	1.1	797	20	AAY05848	Banana ripening fr
	435	8	1.1	801	22	AAM40858	Human polypeptide
	436	8	1.1	802	22	ABB71892	Drosophila melanog
	437	8	1.1	805	22	ABB61739	Drosophila melanog
c	438	8	1.1	805	22	ABB62304	Drosophila melanog
	439	8	1.1	813	22	ABB64239	Drosophila melanog
	440	8	1.1	859	20	AAY00991	Human ATP-dependen
	441	8	1.1	859	22	AAM38810	Human polypeptide
	442	8	1.1	859	22	AAG93268	Human protein HP02
	443	8	1.1	859	22	AAB37463	Human ATP-dependen
	444	8	1.1	872	22	ABG07781	Novel human diagno
	445	8	1.1	887	22	AAM40596	Human polypeptide
	446	8	1.1	892	22	ABG30372	Novel human diagno
c	447	8	1.1	926	23	ABB97572	Novel human protei
c	448	8	1.1	951	22	AAE06694	Arabidopsis thalia
	449	8	1.1	965	21	AAG53461	Arabidopsis thalia
	450	8	1.1	971	21	AAG53460	Arabidopsis thalia
	451	8	1.1	1089	21	AAG53459	Arabidopsis thalia
c	452	8	1.1	1096	22	ABB60222	Drosophila melanog
c	453	8	1.1	1108	22	ABB63413	Drosophila melanog
c	454	8	1.1	1164	22	ABB57802	Drosophila melanog
c	455	8	1.1	1338	22	ABB70361	Drosophila melanog
c	456	8	1.1	1368	22	ABB65408	Drosophila melanog
c	457	8	1.1	1783	22	ABB63930	Drosophila melanog
c	458	8	1.1	2090	22	ABB64682	Drosophila melanog
c	459	8	1.1	2153	22	AAU33195	Novel human secret
	460	8	1.1	2175	22	ABB65698	Drosophila melanog
c	461	8	1.1	2306	23	AAU11817	Cancer and neuroge
c	462	8	1.1	2352	23	AAU11816	Cancer and neuroge
	463	7	1.0	10	22	AAG95019	Human complementar
	464	7	1.0	10	22	AAG86281	Saccharomyces cere
c	465	7	1.0	14	20	AAW67833	Human secreted pro
c	466	7	1.0	17	21	AAY95231	Humanised antibody
	467	7	1.0	18	17	AAW05474	SH3-binding peptid
	468	7	1.0	18	22	ABB40992	Peptide #8498 enco
	469	7	1.0	18	22	ABB25090	Protein #7089 enco
	470	7	1.0	18	22	AAM61852	Human brain expres
	471	7	1.0	18	22	AAM74652	Human bone marrow
	472	7	1.0	18	22	AAM34767	Peptide #8804 enco
	473	7	1.0	18	22	AAB61207	Human INTERCEPT 29
	474	7	1.0	18	23	ABG44475	Human peptide enco
c	475	7	1.0	19	20	AAY11541	Human 5' EST secre
c	476	7	1.0	20	21	AAB45163	Human secreted pro
c	477	7	1.0	20	22	AAB61461	TANGO 300 signal p
c	478	7	1.0	21	21	AAB34122	Human secreted pro
c	479	7	1.0	22	20	AAY31186	araB promoter and
c	480	7	1.0	22	20	AAY36382	Fragment of human
	481	7	1.0	22	20	AAY01199	Polypeptide fragme
c	482	7	1.0	28	22	ABB27725	Human peptide #376
c	483	7	1.0	28	22	ABB32896	Peptide #402 encod
c	484	7	1.0	28	22	ABB18376	Protein #375 encod
c	485	7	1.0	28	22	AAM53698	Human brain expres
c	486	7	1.0	28	22	AAM66081	Human bone marrow
c	487	7	1.0	28	22	AAM13951	Peptide #385 encod
c	488	7	1.0	28	22	AAM26357	Peptide #394 encod
c	489	7	1.0	28	22	AAM01694	Peptide #376 encod
c	490	7	1.0	28	23	ABG35730	Human peptide enco
	491	7	1.0	30	23	ABP41456	Human ovarian anti
	492	7	1.0	31	21	AAB51946	Human secreted pro
c	493	7	1.0	31	22	AAM05988	Peptide #4670 enco
	494	7	1.0	33	21	AAB39001	Human secreted pep
c	495	7	1.0	33	22	AAB65096	Gene #23 associate
c	496	7	1.0	34	22	ABB28682	Peptide #1333 enco
c	497	7	1.0	34	22	ABB33867	Peptide #1373 enco
	498	7	1.0	34	22	ABB40068	Peptide #7574 enco
c	499	7	1.0	34	22	ABB19307	Protein #1306 enco
	500	7	1.0	34	22	ABB24561	Protein #6560 enco
c	501	7	1.0	34	22	AAM54631	Human brain expres
	502	7	1.0	34	22	AAM60824	Human brain expres
c	503	7	1.0	34	22	AAM67037	Human bone marrow
	504	7	1.0	34	22	AAM73504	Human bone marrow
c	505	7	1.0	34	22	AAM14898	Peptide #1332 enco
	506	7	1.0	34	22	AAM19993	Peptide #6427 enco

c	507	7	1.0	34	22	AAM27327	Peptide #1364 enco
	508	7	1.0	34	22	AAM33700	Peptide #7737 enco
c	509	7	1.0	34	22	AAM02622	Peptide #1304 enco
c	510	7	1.0	34	23	ABG36694	Human peptide enco
	511	7	1.0	34	23	ABG43364	Human peptide enco
	512	7	1.0	35	22	AAU27369	Novel bone marrow
c	513	7	1.0	36	20	AAY12142	Human 5' EST secre
c	514	7	1.0	37	22	ABB44142	Peptide #11648 enc
c	515	7	1.0	37	22	ABB27026	Protein #9025 enco
c	516	7	1.0	37	22	AAM77874	Human bone marrow
c	517	7	1.0	37	22	AAM87275	Human immune/haema
c	518	7	1.0	37	22	AAM21771	Peptide #8205 enco
c	519	7	1.0	37	22	AAM38095	Peptide #12132 enc
c	520	7	1.0	37	23	ABG46898	Human peptide enco
c	521	7	1.0	38	22	AAO04430	Human polypeptide
c	522	7	1.0	39	22	AAM91327	Human immune/haema
	523	7	1.0	41	10	AAP90948	Peptide 242. Foot
c	524	7	1.0	41	20	AAY11382	Human 5' EST secre
c	525	7	1.0	41	22	ABB31260	Peptide #3911 enco
c	526	7	1.0	41	22	ABB36468	Peptide #3974 enco
c	527	7	1.0	41	22	ABB21811	Protein #3810 enco
c	528	7	1.0	41	22	AAM57230	Human brain expres
c	529	7	1.0	41	22	AAM69632	Human bone marrow
c	530	7	1.0	41	22	AAO08507	Human polypeptide
c	531	7	1.0	41	22	AAM17448	Peptide #3882 enco
c	532	7	1.0	41	22	AAM29965	Peptide #4002 enco
c	533	7	1.0	41	22	AAM05125	Peptide #3807 enco
c	534	7	1.0	41	23	ABG39254	Human peptide enco
	535	7	1.0	42	20	AAY11632	Human 5' EST secre
c	536	7	1.0	43	19	AAW75128	Human secreted pro
	537	7	1.0	43	22	ABB16861	Human nervous syst
c	538	7	1.0	44	20	AAY19565	Amino acid sequenc
c	539	7	1.0	45	22	AAM92541	Human digestive sy
c	540	7	1.0	45	22	AAM38587	Human colorectal c
c	541	7	1.0	47	21	AAG53805	Arabidopsis thalia
c	542	7	1.0	47	21	AAG55227	Arabidopsis thalia
c	543	7	1.0	47	22	AAE06103	Human gene 63 enco
c	544	7	1.0	47	23	ABG33925	Human secreted pro
c	545	7	1.0	48	18	AAW27289	Human M27-2 secret
c	546	7	1.0	48	18	AAW36947	Protein encoded by
c	547	7	1.0	48	19	AAW74587	Amino acid sequenc
c	548	7	1.0	48	19	AAW48799	C-terminal fragmen
c	549	7	1.0	48	19	AAW68145	Human AChE splice
c	550	7	1.0	48	21	AAG13632	Arabidopsis thalia
c	551	7	1.0	48	21	AAG36558	Arabidopsis thalia
c	552	7	1.0	48	21	AAY87126	Human secreted pro
	553	7	1.0	48	22	AAM82560	Human immune/haema
c	554	7	1.0	49	22	AAG74443	Human colon cancer
	555	7	1.0	50	23	ABP31643	Human ORF616 prote
c	556	7	1.0	50	23	ABP04572	Human ORFX protein
c	557	7	1.0	51	22	AAU43078	Propionibacterium
c	558	7	1.0	52	22	ABG27886	Novel human diagno
c	559	7	1.0	52	23	ABP31800	Human ORF773 prote
	560	7	1.0	53	20	AAY14425	Human secreted pro
c	561	7	1.0	53	22	AAU30376	Novel human secret
c	562	7	1.0	54	21	AAG04510	Arabidopsis thalia
c	563	7	1.0	54	22	AAU51558	Propionibacterium
c	564	7	1.0	54	22	AAO05053	Human polypeptide
c	565	7	1.0	55	20	AAY12547	Human 5' EST secre
c	566	7	1.0	56	17	AAR91275	Eimeria gametocyte
c	567	7	1.0	56	20	AAY27192	Partial amino acid
c	568	7	1.0	56	20	AAW85728	Novel protein (Clo
c	569	7	1.0	56	21	AAY87299	Human signal pepti
	570	7	1.0	56	22	AAO09668	Human polypeptide
c	571	7	1.0	57	20	AAY12935	Amino acid sequenc
	572	7	1.0	57	22	AAU45783	Propionibacterium
	573	7	1.0	57	22	AAM82726	Human immune/haema
c	574	7	1.0	58	22	AAU27459	Novel bone marrow
c	575	7	1.0	58	22	AAO08489	Human polypeptide
c	576	7	1.0	60	19	AAW54440	Mouse novel secret
c	577	7	1.0	60	21	AAB10247	Murine fetal thymu
c	578	7	1.0	60	22	ABB41068	Peptide #8574 enco
c	579	7	1.0	60	22	ABB25133	Protein #7132 enco
c	580	7	1.0	60	22	AAM61926	Human brain expres
c	581	7	1.0	60	22	AAM74728	Human bone marrow
c	582	7	1.0	60	22	AAM20394	Peptide #6828 enco
c	583	7	1.0	60	22	AAM34845	Peptide #8882 enco
c	584	7	1.0	60	23	ABG44532	Human peptide enco
c	585	7	1.0	61	20	AAY02965	Fragment of human
c	586	7	1.0	62	22	ABB16236	Human nervous syst
c	587	7	1.0	62	22	AAM90903	Human immune/haema
c	588	7	1.0	63	21	AAG59384	Arabidopsis thalia
	589	7	1.0	63	22	AAO07988	Human polypeptide
c	590	7	1.0	64	22	ABG29218	Novel human diagno
c	591	7	1.0	64	22	AAG98737	Human cell death p
c	592	7	1.0	65	22	ABG06150	Novel human diagno
	593	7	1.0	65	22	AAB61454	Metallothionein do
c	594	7	1.0	66	21	AAG44141	Arabidopsis thalia
c	595	7	1.0	67	22	AAO07676	Human polypeptide
c	596	7	1.0	68	21	AAB34761	Human secreted pro
c	597	7	1.0	68	22	ABG06151	Novel human diagno
c	598	7	1.0	68	22	ABB17440	Human nervous syst
c	599	7	1.0	68	23	ABP01087	Human ORFX protein
c	600	7	1.0	69	21	AAG16529	Arabidopsis thalia
c	601	7	1.0	69	21	AAG54908	Arabidopsis thalia
c	602	7	1.0	70	23	ABP34865	Human ORF3838 prot
c	603	7	1.0	71	22	AAO08474	Human polypeptide
c	604	7	1.0	71	23	AAU99186	Target molecule hu
c	605	7	1.0	72	20	AAW30658	Human secreted pro
	606	7	1.0	72	21	AAG18831	Zea mays protein f
c	607	7	1.0	72	22	AAU54276	Propionibacterium
	608	7	1.0	72	22	ABG00628	Novel human diagno
c	609	7	1.0	72	22	ABG05186	Novel human diagno
c	610	7	1.0	72	22	AAM84634	Human immune/haema
c	611	7	1.0	72	22	AAM23545	Human EST encoded
c	612	7	1.0	73	20	AAY19450	Amino acid sequenc
c	613	7	1.0	73	21	AAG51998	Arabidopsis thalia
c	614	7	1.0	73	22	ABB68835	Drosophila melanog
	615	7	1.0	73	22	AAU62145	Propionibacterium
c	616	7	1.0	74	21	AAG57266	Arabidopsis thalia
c	617	7	1.0	74	21	AAG00565	Human secreted pro
c	618	7	1.0	75	21	AAG30248	Arabidopsis thalia
c	619	7	1.0	75	22	AAU42451	Propionibacterium
c	620	7	1.0	75	22	AAM89249	Human immune/haema
c	621	7	1.0	75	22	AAG98736	Human cell death p
	622	7	1.0	75	23	ABP31124	Human ORF97 protei
	623	7	1.0	76	22	AAU51607	Propionibacterium
	624	7	1.0	76	22	AAU69856	Human prostate cDN
	625	7	1.0	76	22	AAU69933	Human prostate cDN
	626	7	1.0	76	22	AAM01211	Human prostate-spe
	627	7	1.0	76	22	AAM01288	P789P splice varia
c	628	7	1.0	76	22	AAB48733	Mouse liver growth
	629	7	1.0	76	22	AAE10192	Human bone marrow
	630	7	1.0	76	23	ABB95316	Human P776P splice
	631	7	1.0	76	23	ABB95393	Human P776P splice
c	632	7	1.0	77	15	AAR62651	Androgen receptor
c	633	7	1.0	77	19	AAW48348	Human breast cance
c	634	7	1.0	77	22	ABG18992	Novel human diagno
c	635	7	1.0	77	22	ABB31623	Peptide #4274 enco
c	636	7	1.0	77	22	ABB36840	Peptide #4346 enco
c	637	7	1.0	77	22	ABB22163	Protein #4162 enco
c	638	7	1.0	77	22	AAM57584	Human brain expres
c	639	7	1.0	77	22	AAM69988	Human bone marrow
c	640	7	1.0	77	22	AAM17819	Peptide #4253 enco
c	641	7	1.0	77	22	AAM30326	Peptide #4363 enco
c	642	7	1.0	77	22	AAM05466	Peptide #4148 enco
c	643	7	1.0	77	23	ABG39617	Human peptide enco
c	644	7	1.0	79	21	AAG57689	Arabidopsis thalia
c	645	7	1.0	79	22	AAU50439	Propionibacterium
c	646	7	1.0	81	21	AAY94429	Human Zsig61 prote
c	647	7	1.0	81	22	AAE06619	Human protein havi
c	648	7	1.0	82	21	AAG28789	Arabidopsis thalia
c	649	7	1.0	82	21	AAG57688	Arabidopsis thalia
c	650	7	1.0	82	22	AAU44072	Propionibacterium
	651	7	1.0	83	22	ABB16626	Human nervous syst
	652	7	1.0	83	22	AAO09340	Human polypeptide

c	653	7	1.0	83	23	ABP31946	Human ORF919 prote
c	654	7	1.0	84	21	AAG00172	Human secreted pro
c	655	7	1.0	85	21	AAG51997	Arabidopsis thalia
	656	7	1.0	85	22	AAU62629	Propionibacterium
c	657	7	1.0	85	22	AAM24409	Human EST encoded
	658	7	1.0	85	22	AAG76182	Human colon cancer
c	659	7	1.0	85	23	AAE23312	Human p54S6K prote
c	660	7	1.0	85	23	AAE23318	Human p85S6K prote
	661	7	1.0	85	23	ABP29271	Streptococcus poly
	662	7	1.0	86	20	AAY11933	Human 5' EST secre
c	663	7	1.0	86	21	AAB32792	Eucalyptus grandis
c	664	7	1.0	86	21	AAG26614	Arabidopsis thalia
c	665	7	1.0	86	22	ABB28185	Human peptide #836
c	666	7	1.0	86	22	ABB33360	Peptide #866 encod
c	667	7	1.0	86	22	ABB18820	Protein #819 encod
c	668	7	1.0	86	22	AAM54146	Human brain expres
c	669	7	1.0	86	22	AAM66539	Human bone marrow
	670	7	1.0	86	22	AAO04805	Human polypeptide
c	671	7	1.0	86	22	AAM14412	Peptide #846 encod
c	672	7	1.0	86	22	AAM26825	Peptide #862 encod
c	673	7	1.0	86	22	AAM02139	Peptide #821 encod
c	674	7	1.0	86	23	ABG36191	Human peptide enco
	675	7	1.0	87	20	AAY37377	Chlamydia trachoma
c	676	7	1.0	87	22	ABG29217	Novel human diagno
c	677	7	1.0	87	22	AAM85584	Human immune/haema
c	678	7	1.0	87	23	AAU96975	Human Mch2 protein
	679	7	1.0	88	19	AAW77753	Pyruvate oxidase p
c	680	7	1.0	89	16	AAR75150	Treeshrew (His5,Tr
	681	7	1.0	89	22	AAU44512	Propionibacterium
	682	7	1.0	89	22	AAU67315	Propionibacterium
c	683	7	1.0	89	22	ABB10828	Human ovarian and/
c	684	7	1.0	89	22	AAM94499	Human reproductive
	685	7	1.0	91	22	AAU41291	Propionibacterium
c	686	7	1.0	91	22	AAU30488	Novel human secret
	687	7	1.0	92	18	AAW20240	H. pylori secreted
c	688	7	1.0	92	21	AAB42825	Human ORFX ORF2589
c	689	7	1.0	92	22	ABG16631	Novel human diagno
c	690	7	1.0	93	19	AAW53893	Fragment of chimer
c	691	7	1.0	93	22	ABG29222	Novel human diagno
c	692	7	1.0	94	22	ABB17182	Human nervous syst
c	693	7	1.0	95	21	AAG16528	Arabidopsis thalia
c	694	7	1.0	95	21	AAG51996	Arabidopsis thalia
c	695	7	1.0	95	22	ABB32410	Peptide #5061 enco
c	696	7	1.0	95	22	AAM70797	Human bone marrow
c	697	7	1.0	95	22	AAM06200	Peptide #4882 enco
c	698	7	1.0	96	21	AAG57913	Arabidopsis thalia
c	699	7	1.0	96	22	ABB68955	Drosophila melanog
c	700	7	1.0	97	21	AAG28788	Arabidopsis thalia
c	701	7	1.0	97	22	ABG29215	Novel human diagno
c	702	7	1.0	97	22	ABB52761	Escherichia coli p
c	703	7	1.0	98	21	AAB25084	Plant SDF encoded
	704	7	1.0	98	21	AAG23146	Arabidopsis thalia
c	705	7	1.0	98	21	AAG26613	Arabidopsis thalia
	706	7	1.0	98	21	AAG60540	Arabidopsis thalia
c	707	7	1.0	98	22	ABB68654	Drosophila melanog
	708	7	1.0	98	22	ABG22290	Novel human diagno
c	709	7	1.0	98	22	ABB28261	Human peptide #912
c	710	7	1.0	98	22	ABB33436	Peptide #942 encod
c	711	7	1.0	98	22	ABB18895	Protein #894 encod
c	712	7	1.0	98	22	AAM54221	Human brain expres
c	713	7	1.0	98	22	AAM66615	Human bone marrow
c	714	7	1.0	98	22	AAM14488	Peptide #922 encod
c	715	7	1.0	98	22	AAM26901	Peptide #938 encod
c	716	7	1.0	98	22	AAM38916	Human polypeptide
c	717	7	1.0	98	22	AAM02215	Peptide #897 encod
c	718	7	1.0	98	23	ABG36267	Human peptide enco
	719	7	1.0	99	22	AAU34037	Staphylococcus aur
c	720	7	1.0	99	22	AAU57871	Propionibacterium
	721	7	1.0	99	22	AAM84606	Human immune/haema
	722	7	1.0	100	18	AAW20974	H. pylori secreted
c	723	7	1.0	100	22	AAU42466	Propionibacterium
c	724	7	1.0	100	22	ABG26182	Novel human diagno
c	725	7	1.0	100	23	ABB89757	Human polypeptide

	726	7	1.0	101	19	AAW53327	Rat Puralpha like
c	727	7	1.0	101	20	AAY36379	Fragment of human
c	728	7	1.0	101	21	AAG57912	Arabidopsis thalia
c	729	7	1.0	101	22	AAO11110	Human polypeptide
c	730	7	1.0	102	20	AAY48446	Human prostate can
c	731	7	1.0	103	22	ABG29245	Novel human diagno
	732	7	1.0	104	21	AAB33103	Pinus radiata tran
c	733	7	1.0	104	22	AAM40702	Human polypeptide
	734	7	1.0	104	23	ABP05030	Human ORFX protein
c	735	7	1.0	105	21	AAB56137	Human secreted pro
	736	7	1.0	105	21	AAY70424	Staphylococcus aur
	737	7	1.0	105	22	AAU36624	Staphylococcus aur
	738	7	1.0	105	22	AAU37167	Staphylococcus aur
c	739	7	1.0	105	22	AAO05314	Human polypeptide
	740	7	1.0	106	21	AAB33907	Human secreted pro
c	741	7	1.0	107	22	AAU54125	Propionibacterium
c	742	7	1.0	109	22	ABB59656	Drosophila melanog
c	743	7	1.0	109	22	AAU61871	Propionibacterium
c	744	7	1.0	109	22	ABG06310	Novel human diagno
c	745	7	1.0	109	22	AAE03259	Human gene 6 encod
	746	7	1.0	109	23	ABP38420	Staphylococcus epi
	747	7	1.0	110	22	AAM83311	Human immune/haema
c	748	7	1.0	111	22	AAB47331	FCTR8. Homo sapie
	749	7	1.0	111	23	ABP09196	Human ORFX protein
c	750	7	1.0	111	23	AAU83135	Novel secreted pro
	751	7	1.0	112	19	AAW60650	Human C-C chemokin
	752	7	1.0	112	21	AAB01453	Primate CTACK prot
	753	7	1.0	112	21	AAY57771	Human CC type chem
c	754	7	1.0	112	22	AAO09750	Human polypeptide
c	755	7	1.0	113	22	AAU61707	Propionibacterium
c	756	7	1.0	113	23	ABP41890	Human ovarian anti
c	757	7	1.0	113	23	ABB08792	Human G-substrate
c	758	7	1.0	114	21	AAG54768	Arabidopsis thalia
c	759	7	1.0	115	22	ABG11703	Novel human diagno
	760	7	1.0	118	22	AAU45132	Propionibacterium
c	761	7	1.0	118	22	AAO07195	Human polypeptide
	762	7	1.0	119	22	AAM87975	Human immune/haema
c	763	7	1.0	120	23	AAE15401	Human gonadotropin
c	764	7	1.0	121	22	AAB94940	Human protein sequ
c	765	7	1.0	122	23	ABP08628	Human ORFX protein
c	766	7	1.0	122	23	AAU83217	Novel secreted pro
	767	7	1.0	123	22	AAM80996	Human haematologic
	768	7	1.0	123	22	AAM81965	Human haematologic
c	769	7	1.0	124	22	AAU65976	Propionibacterium
	770	7	1.0	124	22	AAO05830	Human polypeptide
c	771	7	1.0	124	22	AAO05832	Human polypeptide
c	772	7	1.0	124	23	ABP02824	Human ORFX protein
c	773	7	1.0	125	20	AAY12263	Human 5' EST secre
	774	7	1.0	125	21	AAG26755	Zea mays protein f
c	775	7	1.0	126	22	AAO01854	Human polypeptide
c	776	7	1.0	127	21	AAG24107	Arabidopsis thalia
c	777	7	1.0	128	23	ABP38036	Staphylococcus epi
	778	7	1.0	129	21	AAG01659	Human secreted pro
	779	7	1.0	131	20	AAY37793	Chlamydia trachoma
c	780	7	1.0	131	21	AAY76147	Human secreted pro
c	781	7	1.0	131	22	ABG14824	Novel human diagno
c	782	7	1.0	132	21	AAG26612	Arabidopsis thalia
	783	7	1.0	133	22	AAO02265	Human polypeptide
	784	7	1.0	133	23	ABP34915	Human helicase-lik
c	785	7	1.0	134	21	AAG00072	Human secreted pro
c	786	7	1.0	135	22	ABB66419	Drosophila melanog
	787	7	1.0	135	22	AAG73999	Human colon cancer
	788	7	1.0	136	22	AAO13335	Human polypeptide
c	789	7	1.0	136	22	AAB36669	Human secretory pr
c	790	7	1.0	137	18	AAW15777	Protein cognate of
c	791	7	1.0	137	21	AAB38331	Human secreted pro
	792	7	1.0	137	22	ABB96293	Human testicular a
	793	7	1.0	137	22	ABG08154	Novel human diagno
c	794	7	1.0	137	22	ABG29219	Novel human diagno
	795	7	1.0	137	22	AAM95762	Human reproductive
	796	7	1.0	137	22	AAO09468	Human polypeptide
c	797	7	1.0	138	22	AAB65092	Gene #23 associate
	798	7	1.0	139	21	AAG57862	Arabidopsis thalia

	799	7	1.0	139	23	ABB90295	Human polypeptide
c	800	7	1.0	140	22	AAO09801	Human polypeptide
c	801	7	1.0	141	21	AAG48936	Arabidopsis thalia
c	802	7	1.0	141	21	AAG54749	Arabidopsis thalia
c	803	7	1.0	141	22	ABG03186	Novel human diagno
c	804	7	1.0	141	22	ABG06237	Novel human diagno
c	805	7	1.0	141	23	AAU83201	Novel secreted pro
c	806	7	1.0	142	21	AAB19730	Human SECX Clone 4
c	807	7	1.0	142	21	AAY53059	Human secreted pro
c	808	7	1.0	142	22	AAU49575	Propionibacterium
c	809	7	1.0	142	22	AAB90601	Human secreted pro
c	810	7	1.0	142	22	AAB90655	Human secreted pro
c	811	7	1.0	142	22	AAB64668	Human secreted pro
c	812	7	1.0	142	23	ABG65429	Human albumin fusi
c	813	7	1.0	142	23	ABB57420	Human secreted pro
c	814	7	1.0	143	21	AAG17023	Arabidopsis thalia
c	815	7	1.0	143	21	AAG51948	Arabidopsis thalia
c	816	7	1.0	143	22	AAG80959	Human nGPCR55. Ho
c	817	7	1.0	144	15	AAR54819	CHEF-1. AAR54819
c	818	7	1.0	144	21	AAG18062	Arabidopsis thalia
c	819	7	1.0	144	22	AAU56906	Propionibacterium
c	820	7	1.0	144	22	AAG71604	Human olfactory re
	821	7	1.0	145	22	AAU61319	Propionibacterium
c	822	7	1.0	146	22	ABB59619	Drosophila melanog
c	823	7	1.0	147	20	AAY59764	Human normal ovari
c	824	7	1.0	147	22	AAB64715	Gene 3 human secre
c	825	7	1.0	147	22	AAB64716	Human secreted pro
c	826	7	1.0	147	23	ABB90343	Human polypeptide
c	827	7	1.0	149	21	AAG18061	Arabidopsis thalia
c	828	7	1.0	150	23	ABB89042	Human polypeptide
c	829	7	1.0	150	23	AAU83189	Novel secreted pro
c	830	7	1.0	150	23	AAM50708	Human secreted cyt
c	831	7	1.0	151	22	AAM91007	Human immune/haema
c	832	7	1.0	152	21	AAG13627	Arabidopsis thalia
c	833	7	1.0	152	22	ABB96477	Human testicular a
c	834	7	1.0	152	22	AAM95946	Human reproductive
c	835	7	1.0	153	22	AAU42621	Propionibacterium
c	836	7	1.0	154	22	ABB69908	Drosophila melanog
c	837	7	1.0	154	22	AAU58397	Propionibacterium
c	838	7	1.0	154	22	ABG03237	Novel human diagno
c	839	7	1.0	155	22	ABB65265	Drosophila melanog
	840	7	1.0	155	22	AAG90300	C glutamicum prote
c	841	7	1.0	155	22	AAG72986	Olfactory receptor
	842	7	1.0	157	21	AAG06178	Arabidopsis thalia
	843	7	1.0	157	21	AAG42670	Arabidopsis thalia
c	844	7	1.0	157	22	AAU30612	Novel human secret
	845	7	1.0	158	21	AAG06177	Arabidopsis thalia
	846	7	1.0	158	21	AAG42669	Arabidopsis thalia
c	847	7	1.0	158	22	ABG06152	Novel human diagno
	848	7	1.0	159	20	AAY35173	Chlamydia pneumoni
c	849	7	1.0	159	21	AAG26455	Arabidopsis thalia
c	850	7	1.0	159	21	AAG37500	Arabidopsis thalia
c	851	7	1.0	159	22	ABG22371	Novel human diagno
	852	7	1.0	160	21	AAG06176	Arabidopsis thalia
	853	7	1.0	160	21	AAG42668	Arabidopsis thalia
c	854	7	1.0	160	22	AAU48455	Propionibacterium
	855	7	1.0	160	22	AAG74164	Human colon cancer
c	856	7	1.0	161	23	AAO15171	Human stresscopin
c	857	7	1.0	161	23	ABB47842	Listeria monocytog
c	858	7	1.0	163	22	AAM79083	Human protein SEQ
c	859	7	1.0	164	21	AAY44213	Rice phytoene synt
c	860	7	1.0	166	22	AAU87565	Novel central nerv
c	861	7	1.0	166	22	ABG27891	Novel human diagno
	862	7	1.0	167	21	AAG23145	Arabidopsis thalia
	863	7	1.0	167	21	AAG60539	Arabidopsis thalia
	864	7	1.0	168	21	AAY58850	Sorghum resistance
c	865	7	1.0	170	16	AAR64977	Rex region gene pr
c	866	7	1.0	170	21	AAG41297	Arabidopsis thalia
c	867	7	1.0	170	22	AAO12089	Human polypeptide
c	868	7	1.0	171	16	AAR84656	STLVpan-p rex prot
	869	7	1.0	173	21	AAG36737	Arabidopsis thalia
	870	7	1.0	173	21	AAG57861	Arabidopsis thalia
	871	7	1.0	173	21	AAY74629	Neisseria meningit

	872	7	1.0	174	21	AAY74627	Neisseria gonorrhe
	873	7	1.0	174	21	AAY74628	Neisseria meningit
c	874	7	1.0	174	22	AAU23789	Novel human enzyme
c	875	7	1.0	176	19	AAW80254	Amino acid sequenc
	876	7	1.0	176	21	AAG08636	Arabidopsis thalia
	877	7	1.0	176	21	AAG12930	Arabidopsis thalia
	878	7	1.0	176	21	AAG34987	Arabidopsis thalia
	879	7	1.0	176	21	AAG36736	Arabidopsis thalia
	880	7	1.0	177	21	AAG08635	Arabidopsis thalia
	881	7	1.0	177	21	AAG34986	Arabidopsis thalia
c	882	7	1.0	177	22	ABB15381	Human nervous syst
	883	7	1.0	177	22	AAE12464	Human translation
	884	7	1.0	177	22	AAB74325	Human translation
c	885	7	1.0	179	22	AAU53900	Propionibacterium
c	886	7	1.0	180	22	AAU87263	Novel central nerv
c	887	7	1.0	180	22	ABB62875	Drosophila melanog
c	888	7	1.0	180	22	AAM99975	Human expressed po
c	889	7	1.0	180	22	ABB10259	Human cDNA SEQ ID
c	890	7	1.0	180	22	AAU18110	Novel human uterin
c	891	7	1.0	180	22	AAU18352	Human endocrine po
c	892	7	1.0	180	22	AAU18668	Renal and cardiova
c	893	7	1.0	180	22	AAU21652	Novel human neopla
c	894	7	1.0	181	22	ABB69088	Drosophila melanog
	895	7	1.0	181	22	ABB70465	Drosophila melanog
	896	7	1.0	182	22	ABB71567	Drosophila melanog
c	897	7	1.0	183	22	AAU87288	Novel central nerv
c	898	7	1.0	183	22	AAM99978	Human expressed po
c	899	7	1.0	183	22	ABB10284	Human cDNA SEQ ID
c	900	7	1.0	183	22	AAU18119	Novel human uterin
c	901	7	1.0	183	22	AAU18365	Human endocrine po
c	902	7	1.0	183	22	AAU18672	Renal and cardiova
c	903	7	1.0	183	22	AAU21662	Novel human neopla
	904	7	1.0	184	21	AAB43402	Human cancer assoc
c	905	7	1.0	184	21	AAG17504	Arabidopsis thalia
c	906	7	1.0	185	20	AAY02369	Polypeptide identi
c	907	7	1.0	188	22	AAE05352	Mouse secreted pro
c	908	7	1.0	188	23	ABB72362	Murine protein iso
c	909	7	1.0	190	20	AAY36922	Amino acid sequenc
	910	7	1.0	191	22	ABB70426	Drosophila melanog
	911	7	1.0	191	22	ABG24965	Novel human diagno
c	912	7	1.0	192	16	AAR69644	Hepatitis C virus
c	913	7	1.0	192	16	AAR69646	Hepatitis C virus
c	914	7	1.0	192	17	AAR89515	Hepatitis C virus
c	915	7	1.0	192	17	AAR89517	Hepatitis C virus
	916	7	1.0	192	21	AAG42187	Arabidopsis thalia
c	917	7	1.0	192	22	AAG72321	Human olfactory re
	918	7	1.0	193	22	AAU62918	Propionibacterium
c	919	7	1.0	193	23	AAE15927	Human 4566767/84-3
c	920	7	1.0	193	23	AAE15928	Mouse 2745735/84-3
c	921	7	1.0	194	13	AAR22383	Antigen tc-23g. E
	922	7	1.0	194	23	ABB06807	Human nGPCR-Seq101
c	923	7	1.0	196	21	AAG17022	Arabidopsis thalia
c	924	7	1.0	196	21	AAG51947	Arabidopsis thalia
	925	7	1.0	196	23	ABP42436	Human ovarian anti
c	926	7	1.0	197	21	AAG08675	Arabidopsis thalia
c	927	7	1.0	197	21	AAG09051	Arabidopsis thalia
c	928	7	1.0	197	21	AAG42759	Arabidopsis thalia
	929	7	1.0	197	22	AAU23666	Novel human enzyme
	930	7	1.0	197	22	AAU04892	Micromonospora eve
	931	7	1.0	198	21	AAY76219	Human secreted pro
	932	7	1.0	198	22	ABB69251	Drosophila melanog
c	933	7	1.0	199	16	AAR88276	Cytomegalovirus pp
c	934	7	1.0	200	22	ABB61325	Drosophila melanog
c	935	7	1.0	201	22	ABB70218	Drosophila melanog
	936	7	1.0	202	20	AAY41309	Human secreted pro
	937	7	1.0	202	21	AAG09042	Arabidopsis thalia
	938	7	1.0	202	21	AAG38651	Arabidopsis thalia
c	939	7	1.0	202	22	ABG28119	Novel human diagno
c	940	7	1.0	203	16	AAR71361	Human truncated MO
c	941	7	1.0	203	21	AAB38447	Fragment of human
	942	7	1.0	203	22	AAU42496	Propionibacterium
	943	7	1.0	204	22	ABB59405	Drosophila melanog
	944	7	1.0	205	22	AAG72897	Human olfactory re

```
   945     7   1.0    205  23  AAE14609            Human microtubule
   946     7   1.0    207  21  AAG34073            Zea mays protein f
   947     7   1.0    207  21  AAG42186            Arabidopsis thalia
   948     7   1.0    207  22  AAU49163            Propionibacterium
   949     7   1.0    210  21  AAG11097            Arabidopsis thalia
   950     7   1.0    210  22  ABG28848            Novel human diagno
   951     7   1.0    212  20  AAY01198            Polypeptide fragme
   952     7   1.0    213  22  ABB67036            Drosophila melanog
c  953     7   1.0    214  20  AAY06303            Mouse pancreatic-d
   954     7   1.0    214  22  ABG12832            Novel human diagno
c  955     7   1.0    216  17  AAR88670            Fusion protein 52/
c  956     7   1.0    216  21  AAG51946            Arabidopsis thalia
c  957     7   1.0    216  22  ABB67796            Drosophila melanog
c  958     7   1.0    217  21  AAG17021            Arabidopsis thalia
   959     7   1.0    218  13  AAR28645            UL26 protease dele
c  960     7   1.0    218  22  ABB66933            Drosophila melanog
c  961     7   1.0    218  22  AAU19191            Human G protein-co
c  962     7   1.0    219  21  AAY75383            Neisseria meningit
c  963     7   1.0    219  21  AAY75384            Neisseria meningit
c  964     7   1.0    219  23  AAE15935            Rat 3395347/77-324
   965     7   1.0    220  21  AAG08365            Arabidopsis thalia
   966     7   1.0    222  22  AAU61742            Propionibacterium
c  967     7   1.0    222  22  ABG29350            Novel human diagno
   968     7   1.0    222  22  AAG72802            Human olfactory re
   969     7   1.0    222  23  AAU72539            Arabidopsis cell c
c  970     7   1.0    223  21  AAG29539            Arabidopsis thalia
c  971     7   1.0    223  23  ABB48116            Listeria monocytog
c  972     7   1.0    224  20  AAY09535            Human S1-3 protein
c  973     7   1.0    224  22  AAG78055            Human zinc finger
c  974     7   1.0    226  23  AAE15934            Human 4502341_JBC_
c  975     7   1.0    227  22  ABB69369            Drosophila melanog
   976     7   1.0    228  22  ABG18908            Novel human diagno
c  977     7   1.0    228  22  AAU05775            Rice invertase inh
c  978     7   1.0    228  22  AAE05346            Mouse EMR2 protein
   979     7   1.0    229  21  AAG08364            Arabidopsis thalia
   980     7   1.0    231  22  ABG26553            Novel human diagno
   981     7   1.0    232  21  AAG21411            Arabidopsis thalia
   982     7   1.0    232  21  AAG60408            Arabidopsis thalia
   983     7   1.0    233  8  AAP71668            Product of Q regio
   984     7   1.0    236  22  ABG05159            Novel human diagno
   985     7   1.0    236  23  ABP26618            Streptococcus poly
   986     7   1.0    237  21  AAG36735            Arabidopsis thalia
   987     7   1.0    237  22  ABB58541            Drosophila melanog
   988     7   1.0    237  22  ABB67675            Drosophila melanog
   989     7   1.0    238  13  AAR28639            UL26 protease dele
c  990     7   1.0    239  21  AAG07732            Arabidopsis thalia
c  991     7   1.0    240  22  AAM90252            Human immune/haema
c  992     7   1.0    240  22  AAE03207            Human gene 6 encod
c  993     7   1.0    240  22  AAE03225            Human gene 6 encod
c  994     7   1.0    240  22  AAB88389            Human membrane or
c  995     7   1.0    240  23  ABG64408            Human albumin fusi
c  996     7   1.0    240  23  ABG64409            Human albumin fusi
c  997     7   1.0    240  23  ABB49440            Listeria monocytog
c  998     7   1.0    241  22  ABB70942            Drosophila melanog
   999     7   1.0    242  21  AAG08363            Arabidopsis thalia
c1000      7   1.0    242  21  AAG08674            Arabidopsis thalia
```

ALIGNMENTS

```
RESULT 1
AAE02417
ID   AAE02417 standard; Protein; 724 AA.
XX
AC   AAE02417;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human vasa protein.
XX
KW   Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;
KW   ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;
KW   colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;
KW   skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;
KW   medulloblastoma; choriocarcinoma; squamous cell carcinoma; leukaemia;
KW   acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;
KW   osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          1..318
FT                   /note= "Highly antigenic"
FT   Region          117..146
FT                   /label= Epitope #1
FT   Region          695..723
FT                   /label= Epitope #2
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
DR   N-PSDB; AAD06354.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 18; Page 53-54; 66pp; English.
XX
CC   The present sequence is human vasa protein that has germ cell specific
CC   expression and is believed to play a determinative role in gonad
CC   development. Germ cells are specialised to produce haploid gametes in
CC   multicellular organisms. Vasa is useful in the diagnosis or treatment
CC   of conditions characterised by its aberrant expression and/or the
CC   presence of its mutant forms. The conditions include cancers such as
CC   biliary tract, brain, breast, colon, ovarian, pancreas, prostate,
CC   colorectal, oral, liver, lung, skin, basocellular, testis, renal,
CC   thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,
CC   melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   724 AA;

Alignment Scores:
Pred. No.:               0             Length:        724
Score:                   724.00        Matches:       724
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      22            Gaps:          0

US-09-714-865-15 (1-2172) x AAE02417 (1-724)

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAspGluAspTrpGluAlaGluIleAsnProHisMetSerSerTyrValProIle 20
```

```
Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 PheGluLysAspArgTyrSerGlyGluAsnGlyAspAsnPheAsnArgThrProAlaSer 40

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerSerGluMetAspAspGlyProSerArgArgAspHisPheMetLysSerGlyPheAla 60

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerGlyArgAsnPheGlyAsnArgAspAlaGlyGluCysAsnLysArgAspAsnThrSer 80

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrMetGlyGlyPheGlyValGlyLysSerPheGlyAsnArgGlyPheSerAsnSerArg 100

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheGluAspGlyAspSerSerGlyPheTrpArgGluSerSerAsnAspCysGluAspAsn 120

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ProThrArgAsnArgGlyPheSerLysArgGlyGlyTyrArgAspGlyAsnAsnSerGlu 140

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AlaSerGlyProTyrArgArgGlyGlyArgGlySerPheArgGlyCysArgGlyGlyPhe 160

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GlyLeuGlySerProAsnAsnAspLeuAspProAspGluCysMetGlnArgThrGlyGly 180

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LeuPheGlySerArgArgProValLeuSerGlyThrGlyAsnGlyAspThrSerGlnSer 200

Qy        601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ArgSerGlySerGlySerGluArgGlyGlyTyrLysGlyLeuAsnGluGluValIleThr 220

Qy        661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyGluSerSerAspThrGln 240

Qy        721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePheAla 260

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSerGly 280

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 HisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeuAsn 300

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIlePro 320

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 IleIleLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 340

Qy       1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaPheLeuLeuProIleLeuAlaHisMetMetHisAspGlyIleThrAlaSerArgPhe 360

Qy       1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuValAsnGln 380

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaValValIleTyr 400

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyGlyThrGlnLeuGlyHisSerIleArgGlnIleValGlnGlyCysAsnIleLeuCys 420

Qy       1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLysGlnIle 440

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlyProGluMet 460

Qy       1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 LysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeuMetPhe 480

Qy       1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SerAlaThrPheProGluGluIleGlnArgLeuAlaAlaGluPheLeuLysSerAsnTyr 500

Qy       1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThrValLeu 520

Qy       1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GlnValGlyGlnPheSerLysArgGluLysLeuValGluIleLeuArgAsnIleGlyAsp 540

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GluArgThrMetValPheValGluThrLysLysLysAlaAspPheThrAlaThrPheLeu 560

Qy       1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 CysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGluArgGlu 580

Qy       1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 GlnAlaLeuGlyAspPheArgPheGlyLysCysProValLeuValAlaThrSerValAla 600

Qy       1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAspLeuProSerThr 620

Qy       1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 IleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAla 640

Qy       1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 IleSerPhePheAspLeuGluSerAspAsnHisLeuAlaGlnProLeuValLysValLeu 660

Qy       1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ThrAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerThrTyrIle 680

Qy       2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 ProGlyPheSerGlySerThrArgGlyAsnValPheAlaSerValAspThrArgLysGly 700

Qy       2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 LysSerThrLeuAsnThrAlaGlyPheSerSerSerArgAlaProAsnProValAspAsp 720

Qy       2161 GAGTCATGGGAT 2172
              ||||||||||||
Db        721 GluSerTrpAsp 724
```

RESULT 2

```
AAU23109
ID   AAU23109 standard; Protein; 135 AA.
XX
AC   AAU23109;
XX
DT   17-DEC-2001  (first entry)
XX
DE   Novel human enzyme polypeptide #195.
XX
KW   Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW   ligase; hyperproliferative disorder; immunodeficiency disorder;
KW   autoimmune disorder; neurological disorder; metabolic disorder;
KW   inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW   blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW   nephrotropic; anticoagulant.
XX
OS   Homo sapiens.
XX
PN   WO200155301-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01239.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
```

```
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465566/50.
DR   N-PSDB; AAS40979.
XX
PT   Novel polypeptides and polynucleotides useful for diagnosing,
PT   preventing, treating neural, immune system, muscular, reproductive,
PT   pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT   diseases  -
XX
PS   Claim 11; SEQ ID No 1105; 1180pp; English.
XX
CC   The present invention relates to the isolation of novel human enzyme
CC   polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC   encoding them. The enzyme polypeptides of the invention may comprise the
CC   functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC   isomerases or ligases. The sequences of the invention are useful in the
CC   diagnosis, treatment, prevention and/or prognosis of a wide range of
CC   disorders including hyperproliferative disorders (e.g. cancer),
CC   immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC   (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC   metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC   (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC   blood-related disorders (e.g. haemophilia), reproductive disorders
CC   (e.g. infertility) and infectious disorders (e.g. Influenza). The
CC   polynucleotides of the invention can also be used in gene therapy.
CC   AAU22915-AAU23814 represent the novel human enzyme polypeptides of the
CC   invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   135 AA;

Alignment Scores:
Pred. No.:              9.91e-111     Length:        135
Score:                  123.00        Matches:       123
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            16.99%        Indels:        0
DB:                     22            Gaps:          0

US-09-714-865-15 (1-2172) x AAU23109 (1-135)

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LysCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluAsnVal 20

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlnHisValIleAsnPheAspLeuProSerThrIleAspGluTyrValHisArgIleGly 40

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGluSerAsp 60

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValProAla 80

Qy       2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 TrpLeuGluGluIleAlaPheSerThrTyrIleProGlyPheSerGlySerThrArgGly 100

Qy       2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 AsnValPheAlaSerValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGlyPhe 120

Qy       2128 TCTTCTTCA 2136
              |||||||||
Db        121 SerSerSer 123

RESULT 3
ABB95898
ID   ABB95898 standard; Protein; 106 AA.
XX
AC   ABB95898;
XX
DT   21-JUN-2002  (first entry)
XX
DE   Human testicular antigen SEQ ID NO: 1282.
XX
KW   Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW   reproductive system disorder; urinary system disorder; gene therapy;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disease; infection; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200155317-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01329.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
```

```
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483232/52.
XX
PT   Nucleic acids encoding 973 human testicular antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating testicular cancer  -
XX
PS   Claim 11; SEQ ID NO 1282; 766pp; English.
XX
CC   The present invention provides the protein and coding sequences of 973
CC   human testicular antigens, and fragments of their genomic sequences. The
CC   sequences can be used in the treatment of cardiovascular, urinary system,
CC   reproductive system, immune, respiratory, neurological and
CC   gastrointestinal disorders, infections, and particularly cancer,
CC   especially testicular cancers. The present sequence is a protein of the
CC   invention.
XX
SQ   Sequence   106 AA;

Alignment Scores:
Pred. No.:                1.09e-72        Length:         106
Score:                    84.00           Matches:        84
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              11.60%          Indels:         0
DB:                       22              Gaps:           0

US-09-714-865-15 (1-2172) x ABB95898 (1-106)

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCG 1944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 GlyArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGluSer 30

Qy       1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 AspAsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValPro 50

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGA 2064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 AlaTrpLeuGluGluIleAlaPheSerThrTyrIleProGlyPheSerGlySerThrArg 70
```

```
Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGG 2124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GlyAsnValPheAlaSerValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGly 90

Qy       2125 TTTTCTTCTTCA 2136
              ||||||||||||
Db         91 PheSerSerSer 94

RESULT 4
AAU23664
ID   AAU23664 standard; Protein; 106 AA.
XX
AC   AAU23664;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human enzyme polypeptide #750.
XX
KW   Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW   ligase; hyperproliferative disorder; immunodeficiency disorder;
KW   autoimmune disorder; neurological disorder; metabolic disorder;
KW   inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW   blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW   nephrotropic; anticoagulant.
XX
OS   Homo sapiens.
XX
PN   WO200155301-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01239.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
```

```
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465566/50.
DR   N-PSDB; AAS41534.
XX
PT   Novel polypeptides and polynucleotides useful for diagnosing,
PT   preventing, treating neural, immune system, muscular, reproductive,
PT   pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT   diseases  -
XX
PS   Claim 11; SEQ ID No 1660; 1180pp; English.
XX
CC   The present invention relates to the isolation of novel human enzyme
CC   polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC   encoding them. The enzyme polypeptides of the invention may comprise the
CC   functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC   isomerases or ligases. The sequences of the invention are useful in the
CC   diagnosis, treatment, prevention and/or prognosis of a wide range of
CC   disorders including hyperproliferative disorders (e.g. cancer),
CC   immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC   (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC   metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC   (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC   blood-related disorders (e.g. haemophilia), reproductive disorders
CC   (e.g. infertility) and infectious disorders (e.g. Influenza). The
CC   polynucleotides of the invention can also be used in gene therapy.
CC   AAU22915-AAU23814 represent the novel human enzyme polypeptides of the
CC   invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   106 AA;

Alignment Scores:
Pred. No.:              1.09e-72      Length:         106
Score:                  84.00         Matches:        84
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            11.60%        Indels:         0
DB:                     22            Gaps:           0

US-09-714-865-15 (1-2172) x AAU23664 (1-106)

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCG 1944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 GlyArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGluSer 30

Qy       1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 AspAsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValPro 50

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGA 2064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 AlaTrpLeuGluGluIleAlaPheSerThrTyrIleProGlyPheSerGlySerThrArg 70

Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGG 2124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GlyAsnValPheAlaSerValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGly 90

Qy       2125 TTTTCTTCTTCA 2136
              ||||||||||||
Db         91 PheSerSerSer 94

RESULT 5
AAM95199
ID   AAM95199 standard; Protein; 106 AA.
XX
AC   AAM95199;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen SEQ ID NO: 3857.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
```

```
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465570/50.
DR   N-PSDB; AAL01169.
XX
PT   Isolated nucleic acid molecule encoding a reproductive system antigen
PT   is used in preventing, treating or ameliorating a medical condition  -
XX
PS   Claim 11; SEQ ID NO 3857; 1297pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human reproductive system related antigens. These can be used
CC   in the prevention and treatment of reproductive system disorders,
CC   including cancer. The present sequence is a protein of the invention.
XX
SQ   Sequence   106 AA;

Alignment Scores:
Pred. No.:              1.09e-72      Length:        106
Score:                  84.00         Matches:       84
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            11.60%        Indels:        0
DB:                     22            Gaps:          0

US-09-714-865-15 (1-2172) x AAM95199 (1-106)

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCG 1944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 GlyArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGluSer 30

Qy       1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 AspAsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValPro 50

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGA 2064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 AlaTrpLeuGluGluIleAlaPheSerThrTyrIleProGlyPheSerGlySerThrArg 70

Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGG 2124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GlyAsnValPheAlaSerValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGly 90
```

```
Qy       2125 TTTTCTTCTTCA 2136
              ||||||||||||
Db         91 PheSerSerSer 94

RESULT 6
AAE02419
ID   AAE02419 standard; Protein; 713 AA.
XX
AC   AAE02419;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Rat vasa protein.
XX
KW   Rat; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Rattus norvegicus.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 56-58; 66pp; English.
XX
CC   The present sequence is rat vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   713 AA;

Alignment Scores:
Pred. No.:              7.88e-71     Length:         713
Score:                  82.00        Matches:        82
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            11.33%       Indels:         0
DB:                     22           Gaps:           0

US-09-714-865-15 (1-2172) x AAE02419 (1-713)

Qy        718 CAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePhe 244

Qy        778 GCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 AlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSer 264

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeu 284

Qy        898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 AsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIle 304

Qy        958 CCTATC 963
              ||||||
Db        305 ProIle 306

RESULT 7
AAE02418
ID   AAE02418 standard; Protein; 722 AA.
XX
AC   AAE02418;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Mouse vasa protein.
XX
KW   Mouse; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Mus musculus.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 54-56; 66pp; English.
XX
CC   The present sequence is mouse vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
```

```
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   722 AA;

Alignment Scores:
Pred. No.:              2.81e-53      Length:         722
Score:                  64.00         Matches:        64
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            8.84%         Indels:         0
DB:                     22            Gaps:           0

US-09-714-865-15 (1-2172) x AAE02418 (1-722)

Qy        718 CAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 GlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePhe 258

Qy        778 GCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 AlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSer 278

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeu 298

Qy        898 AATAACAACATT 909
              ||||||||||||
Db        299 AsnAsnAsnIle 302

RESULT 8
AAE02423
ID   AAE02423 standard; peptide; 30 AA.
XX
AC   AAE02423;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Epitope #1 of human vasa protein, used to generate vasa antibodies.
XX
KW   Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;
KW   ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;
KW   colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;
KW   skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;
KW   medulloblastoma; choriocarcinoma; squamous cell carcinoma; leukaemia;
KW   acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;
KW   osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Homo sapiens.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 19; Page 36; 66pp; English.
XX
CC   The present sequence is epitope of human vasa protein that has germ cell
CC   specific expression and is believed to play a determinative role in gonad
CC   Epitope is used in the generation of anti-human vasa antibodies.
CC   development. Germ cells are specialised to produce haploid gametes in
CC   multicellular organisms. Vasa is useful in the diagnosis or treatment
CC   of conditions characterised by its aberrant expression and/or the
CC   presence of its mutant forms. The conditions include cancers such as
CC   biliary tract, brain, breast, colon, ovarian, pancreas, prostate,
CC   colorectal, oral, liver, lung, skin, basocellular, testis, renal,
CC   thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,
CC   melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   30 AA;

Alignment Scores:
Pred. No.:              5.75e-20      Length:         30
Score:                  30.00         Matches:        30
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            4.14%         Indels:         0
DB:                     22            Gaps:           0

US-09-714-865-15 (1-2172) x AAE02423 (1-30)

Qy        349 TGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGA 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CysGluAspAsnProThrArgAsnArgGlyPheSerLysArgGlyGlyTyrArgAspGly 20

Qy        409 AATAATTCAGAAGCTTCAGGGCCATACAGA 438
              ||||||||||||||||||||||||||||||
Db         21 AsnAsnSerGluAlaSerGlyProTyrArg 30

RESULT 9
AAE02424
ID   AAE02424 standard; peptide; 29 AA.
XX
AC   AAE02424;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Epitope #2 of human vasa protein, used to generate vasa antibodies.
XX
KW   Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;
KW   ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;
KW   colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;
KW   skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;
KW   medulloblastoma; choriocarcinoma; squamous cell carcinoma; leukaemia;
KW   acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;
KW   osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Homo sapiens.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
```

```
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 19; Page 36; 66pp; English.
XX
CC   The present sequence is epitope of human vasa protein that has germ cell
CC   specific expression and is believed to play a determinative role in gonad
CC   Epitope is used in the generation of anti-human vasa antibodies.
CC   development. Germ cells are specialised to produce haploid gametes in
CC   multicellular organisms. Vasa is useful in the diagnosis or treatment
CC   of conditions characterised by its aberrant expression and/or the
CC   presence of its mutant forms. The conditions include cancers such as
CC   biliary tract, brain, breast, colon, ovarian, pancreas, prostate,
CC   colorectal, oral, liver, lung, skin, basocellular, testis, renal,
CC   thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,
CC   melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   29 AA;

Alignment Scores:
Pred. No.:              5.45e-19     Length:         29
Score:                  29.00        Matches:        29
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            4.01%        Indels:         0
DB:                     22           Gaps:           0

US-09-714-865-15 (1-2172) x AAE02424 (1-29)

Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGlyPheSerSerSerArgAla 20

Qy       2143 CCCAATCCAGTAGATGATGAGTCATGG 2169
              |||||||||||||||||||||||||||
Db         21 ProAsnProValAspAspGluSerTrp 29

RESULT 10
AAE02420
ID   AAE02420 standard; Protein; 700 AA.
XX
AC   AAE02420;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Frog vasa protein.
XX
KW   Frog; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Xenopus laevis.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 58-59; 66pp; English.
XX
CC   The present sequence is frog vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   700 AA;

Alignment Scores:
Pred. No.:              2.86e-14     Length:         700
Score:                  24.00        Matches:        24
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            3.31%        Indels:         0
DB:                     22           Gaps:           0

US-09-714-865-15 (1-2172) x AAE02420 (1-700)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 329

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        330 LeuProIleLeu 333

RESULT 11
AAE02421
ID   AAE02421 standard; Protein; 700 AA.
XX
AC   AAE02421;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Danio reio vasa protein.
XX
```

```
KW   Vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial.
XX
OS   Danio reio.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 59-61; 66pp; English.
XX
CC   The present sequence is Danio reio vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   700 AA;

Alignment Scores:
Pred. No.:               2.86e-14     Length:         700
Score:                   24.00        Matches:        24
Percent Similarity:      100.00%      Conservative:   0
Best Local Similarity:   100.00%      Mismatches:     0
Query Match:             3.31%        Indels:         0
DB:                      22           Gaps:           0

US-09-714-865-15 (1-2172) x AAE02421 (1-700)

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 316

Qy       1030 CTACCAATTTTG 1041
              ||||||||||||
Db        317 LeuProIleLeu 320

RESULT 12
ABB59954
ID   ABB59954 standard; Protein; 661 AA.
XX
AC   ABB59954;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 6654.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL04057.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 6654; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   661 AA;

Alignment Scores:
Pred. No.:               2.72e-13     Length:         661
Score:                   23.00        Matches:        23
Percent Similarity:      100.00%      Conservative:   0
Best Local Similarity:   100.00%      Mismatches:     0
Query Match:             3.18%        Indels:         0
DB:                      22           Gaps:           0

US-09-714-865-15 (1-2172) x ABB59954 (1-661)

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeu 301

Qy       1033 CCAATTTTG 1041
              |||||||||
Db        302 ProIleLeu 304

RESULT 13
AAE02422
ID   AAE02422 standard; Protein; 661 AA.
XX
AC   AAE02422;
XX
DT   10-AUG-2001  (first entry)
XX
```

```
DE   Fruit fly vasa protein.
XX
KW   Vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KW   pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KW   thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KW   melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KW   choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KW   myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KW   Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial; fruit fly.
XX
OS   Drosophila melanogaster.
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 4; Page 61-62; 66pp; English.
XX
CC   The present sequence is fruit fly vasa protein. Vasa is useful in the
CC   diagnosis or treatment of conditions characterised by its aberrant
CC   expression and/or the presence of its mutant forms. The conditions
CC   include cancers such as biliary tract, brain, breast, colon, ovarian,
CC   pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC   testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC   lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence   661 AA;

Alignment Scores:
Pred. No.:              2.72e-13      Length:        661
Score:                  23.00         Matches:       23
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.18%         Indels:        0
DB:                     22            Gaps:          0

US-09-714-865-15 (1-2172) x AAE02422 (1-661)

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeu 301

Qy       1033 CCAATTTTG 1041
              |||||||||
Db        302 ProIleLeu 304

RESULT 14
AAB38176
ID   AAB38176 standard; Protein; 112 AA.
XX
AC   AAB38176;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Gene 6 human secreted protein homologous amino acid sequence #115.
XX
KW   Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
KW   hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW   cerebrovascular disorder; nervous system disorder; ocular disorder;
KW   wound healing; skin aging; food additive; preservative.
XX
OS   Homo sapiens.
XX
PN   WO200058468-A2.
XX
PD   05-OCT-2000.
XX
PF   22-MAR-2000; 2000WO-US07526.
XX
PR   26-MAR-1999;   99US-0126600.
PR   22-DEC-1999;   99US-0171550.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-611713/58.
XX
PT   Nucleic acids encoding human secreted proteins, used to prevent, treat,
PT   ameliorate, or diagnose conditions such as autoimmune disorders, skin
PT   disorders and cancer -
XX
PS   Disclosure; Page 363; 374pp; English.
XX
CC   The polynucleotide sequences given in AAC69399 to AAC69445 encode the
CC   human secreted proteins given in AAB38119 to AAB38165. AAB38166 to
CC   AAB38201 represent human secreted polypeptide sequences and proteins
CC   homologous to them, which are given in the exemplification of the present
CC   invention. Human secreted proteins have activities based on the tissues
CC   and cells the genes are expressed in. Example of activities include:
CC   immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
CC   cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
CC   neuroprotective; antibacterial; virucide; fungicide; and
CC   ophthalmological. The polynucleotides and polypeptides can be are used to
CC   prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC   in diagnosing a pathological condition or susceptibility to a
CC   pathological condition. Disorders which are diagnosed or treated include
CC   autoimmune diseases, hyperproliferative disorders, cardiovascular
CC   disorders, cerebrovascular disorders, angiogenesis, nervous system
CC   disorders, infections caused by bacteria, viruses and fungi and ocular
CC   disorders. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities. AAC69390 to AAC69398 and
CC   AAB38118 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   112 AA;

Alignment Scores:
Pred. No.:              3.14e-12      Length:        112
Score:                  22.00         Matches:       22
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.04%         Indels:        0
```

```
DB:                     21          Gaps:        0

US-09-714-865-15 (1-2172) x AAB38176 (1-112)

Qy         976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          71 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 90

Qy        1036 ATTTTG 1041
               ||||||
Db          91 IleLeu 92

RESULT 15
AAB38177
ID   AAB38177 standard; Protein; 112 AA.
XX
AC   AAB38177;
XX
DT   30-JAN-2001  (first entry)
XX
DE   Human secreted protein sequence encoded by gene 6 SEQ ID NO:116.
XX
KW   Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
KW   hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW   cerebrovascular disorder; nervous system disorder; ocular disorder;
KW   wound healing; skin aging; food additive; preservative.
XX
OS   Homo sapiens.
XX
PN   WO200058468-A2.
XX
PD   05-OCT-2000.
XX
PF   22-MAR-2000; 2000WO-US07526.
XX
PR   26-MAR-1999;   99US-0126600.
PR   22-DEC-1999;   99US-0171550.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-611713/58.
XX
PT   Nucleic acids encoding human secreted proteins, used to prevent, treat,
PT   ameliorate, or diagnose conditions such as autoimmune disorders, skin
PT   disorders and cancer -
XX
PS   Disclosure; Page 363-364; 374pp; English.
XX
CC   The polynucleotide sequences given in AAC69399 to AAC69445 encode the
CC   human secreted proteins given in AAB38119 to AAB38165. AAB38166 to
CC   AAB38201 represent human secreted polypeptide sequences and proteins
CC   homologous to them, which are given in the exemplification of the present
CC   invention. Human secreted proteins have activities based on the tissues
CC   and cells the genes are expressed in. Example of activities include:
CC   immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
CC   cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
CC   neuroprotective; antibacterial; virucide; fungicide; and
CC   ophthalmological. The polynucleotides and polypeptides can be are used to
CC   prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC   in diagnosing a pathological condition or susceptibility to a
CC   pathological condition. Disorders which are diagnosed or treated include
CC   autoimmune diseases, hyperproliferative disorders, cardiovascular
CC   disorders, cerebrovascular disorders, angiogenesis, nervous system
CC   disorders, infections caused by bacteria, viruses and fungi and ocular
CC   disorders. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities. AAC69390 to AAC69398 and
CC   AAB38118 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   112 AA;

Alignment Scores:
Pred. No.:              3.14e-12      Length:         112
Score:                  22.00         Matches:        22
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            3.04%         Indels:         0
DB:                     21            Gaps:           0

US-09-714-865-15 (1-2172) x AAB38177 (1-112)

Qy         976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          71 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 90

Qy        1036 ATTTTG 1041
               ||||||
Db          91 IleLeu 92

Search completed: June 10, 2003, 17:04:54
Job time : 82.5 secs
```

```
                GenCore version 5.1.6
      Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         June 10, 2003, 11:55:35 ; Search time 2772 Seconds
                                          (without alignments)
                                          12689.970 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  2172
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched:       16154066 seqs, 8097743376 residues

Word size :     0

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1806	83.1	2523	11	BC030638	BC030638 Homo sapi
	2	616	28.4	682	12	BG718234	BG718234 602696209
	3	616	28.4	693	13	BI560849	BI560849 603254018
	4	612	28.2	798	12	BG717110	BG717110 602689191
	5	535	24.6	792	14	BQ429126	BQ429126 AGENCOURT
	6	509	23.4	747	12	BG719449	BG719449 602690172
	7	480	22.1	657	12	BG717812	BG717812 602693829
	8	469	21.6	651	12	BG717354	BG717354 602689688
	9	441	20.3	822	13	BI464267	BI464267 603203379
	10	440	20.3	846	12	BG718634	BG718634 602696739
	11	430	19.8	861	13	BI461393	BI461393 603206861
	12	422	19.4	830	13	BI462207	BI462207 603205520
	13	413	19.0	699	12	BG771822	BG771822 602720504
	14	412	19.0	743	13	BI463502	BI463502 603203596
	15	392	18.0	841	13	BI459115	BI459115 603199485
	16	381	17.5	884	13	BI560499	BI560499 603254476
	17	361	16.6	918	13	BI561001	BI561001 603254391
	18	337	15.5	810	13	BI463846	BI463846 603202572
	19	328	15.1	758	13	BI559479	BI559479 603252992
	20	321	14.8	482	9	AL042306	AL042306 DKFZp434M
c	21	277	12.8	555	9	AA399611	AA399611 zt39a07.s
	22	267	12.3	847	13	BI562624	BI562624 603256559
c	23	246	11.3	555	9	AI217144	AI217144 qf47d11.x
	24	246	11.3	917	13	BI463447	BI463447 603204483
	25	233	10.7	335	9	AA383535	AA383535 EST96928
	26	219	10.1	491	9	AA398976	AA398976 zt93a07.r
	27	204	9.4	700	12	BG718861	BG718861 602699337
c	28	202	9.3	347	9	AI953070	AI953070 wq49h07.x
	29	161	7.4	604	13	BI460628	BI460628 603201271
c	30	132	6.1	300	9	AI337133	AI337133 qx83b04.x
c	31	132	6.1	300	9	AI654417	AI654417 wb31f04.x
c	32	132	6.1	300	10	BE550307	BE550307 7a21g08.x
c	33	132	6.1	300	10	BE550448	BE550448 7a25e08.x
c	34	114	5.2	360	10	AW449170	AW449170 UI-H-BI3-
c	35	108	5.0	108	9	AI969018	AI969018 wq68d02.x
c	36	64	2.9	371	9	AA758412	AA758412 ah66g05.s
c	37	63	2.9	469	9	AI025074	AI025074 ov40a04.x
	38	58	2.7	2759	11	AK014844	AK014844 Mus muscu
c	39	53	2.4	374	17	B63809	B63809 RPCI11-3J20
c	40	49	2.3	103	9	AA400066	AA400066 zu62a08.s
	41	44	2.0	546	13	BM539751	BM539751 hb13g02.g
c	42	42	1.9	665	17	AZ648249	AZ648249 1M0517K12
c	43	41	1.9	527	17	AZ806430	AZ806430 2M0068P20
c	44	41	1.9	589	17	AQ932766	AQ932766 RPCI-23-2
	45	39	1.8	79	9	AA401568	AA401568 zu62a08.r
c	46	39	1.8	100	9	AA862553	AA862553 oh44b08.s
	47	36	1.7	772	13	BI830532	BI830532 603073238
	48	34	1.6	637	13	BI461238	BI461238 603206562
	49	33	1.5	670	14	BQ443738	BQ443738 UI-M-EW0-
	50	28	1.3	420	17	AQ543635	AQ543635 RPCI-11-3
c	51	26	1.2	430	9	AI843771	AI843771 UI-M-AN1-
c	52	25	1.2	737	17	AG123731	AG123731 Pan trogl
c	53	23	1.1	141	12	BE984091	BE984091 UI-M-CG0p
	54	23	1.1	346	10	BE192466	BE192466 EST331231
	55	23	1.1	355	9	AA087515	AA087515 mm88a03.r
	56	23	1.1	435	10	BE227079	BE227079 us76e10.y
c	57	23	1.1	441	9	AA517647	AA517647 vh81d02.r
	58	23	1.1	466	13	BI190619	BI190619 i3a01fs.r
	59	23	1.1	497	12	BF150205	BF150205 uy83a07.y
	60	23	1.1	518	13	BM125854	BM125854 id76b04.y
	61	23	1.1	555	17	AZ401250	AZ401250 1M0167D24
	62	23	1.1	587	9	AA896298	AA896298 vy13a04.r
	63	23	1.1	600	12	BG804318	BG804318 0254-22 M
	64	23	1.1	600	12	BG808188	BG808188 2082-77 M
	65	23	1.1	610	10	BE382075	BE382075 601272235
	66	23	1.1	621	9	AA098277	AA098277 mn94c10.r
	67	23	1.1	666	12	BF452394	BF452394 uz89c02.y
	68	23	1.1	696	14	BQ265507	BQ265507 NISC_ff06
	69	23	1.1	714	12	BF166960	BF166960 601773715
	70	23	1.1	734	12	BF166901	BF166901 601774132
	71	23	1.1	753	13	BI689003	BI689003 603313488
	72	23	1.1	762	13	BG974272	BG974272 602844017
	73	23	1.1	790	13	BI078164	BI078164 602872482
	74	23	1.1	824	12	BE916016	BE916016 601666315
	75	23	1.1	847	12	BG519330	BG519330 602577585
	76	23	1.1	855	12	BF607262	BF607262 MY1_00020
	77	23	1.1	864	12	BG293320	BG293320 602390708
	78	23	1.1	891	9	AI324427	AI324427 mm88a03.y
	79	23	1.1	906	14	BQ964084	BQ964084 AGENCOURT

	80	23	1.1	941	14	BQ886382	BQ886382 AGENCOURT
	81	23	1.1	959	14	BQ714521	BQ714521 AGENCOURT
	82	23	1.1	965	14	BQ714287	BQ714287 AGENCOURT
	83	23	1.1	969	14	BQ928214	BQ928214 AGENCOURT
	84	23	1.1	1099	12	BE913621	BE913621 601667566
	85	22	1.0	604	17	AZ448672	AZ448672 1M0246D03
	86	22	1.0	635	12	BG448350	BG448350 NF070H07E
	87	22	1.0	727	12	BG645827	BG645827 EST507446
	88	21	1.0	255	9	AV125950	AV125950 AV125950
c	89	21	1.0	495	12	BG512399	BG512399 dad20b08.
	90	21	1.0	530	17	AQ589388	AQ589388 CITBI-E1-
c	91	21	1.0	588	14	BQ780602	BQ780602 UI-R-FF0-
	92	21	1.0	693	17	AQ854670	AQ854670 CpG2132A
	93	21	1.0	780	10	AW153705	AW153705 fi24g05.y
	94	21	1.0	860	9	AI049281	AI049281 uc85f11.y
	95	21	1.0	999	14	BQ891158	BQ891158 AGENCOURT
c	96	21	1.0	1007	17	CNS03BC3	AL236316 Tetraodon
c	97	20	0.9	190	14	BQ756631	BQ756631 EBem09_SQ
	98	20	0.9	247	9	AV274771	AV274771 AV274771
c	99	20	0.9	247	10	BB152670	BB152670 BB152670
c	100	20	0.9	282	9	AI795264	AI795264 RRAMCA966
c	101	20	0.9	300	9	AJ464817	AJ464817 AJ464817
c	102	20	0.9	300	9	AJ472851	AJ472851 AJ472851
c	103	20	0.9	327	14	BQ458661	BQ458661 HA04J08r
c	104	20	0.9	357	13	BM371273	BM371273 EBma08_SQ
c	105	20	0.9	360	9	AJ460815	AJ460815 AJ460815
c	106	20	0.9	360	9	AJ464819	AJ464819 AJ464819
c	107	20	0.9	360	9	AJ481941	AJ481941 AJ481941
	108	20	0.9	360	14	C39342	C39342 C39342 Yuji
c	109	20	0.9	374	13	BM096712	BM096712 EBma07_SQ
	110	20	0.9	382	12	BG405848	BG405848 sac38f03.
c	111	20	0.9	394	13	BM373875	BM373875 EBma03_SQ
c	112	20	0.9	397	13	BM444825	BM444825 EBem09_SQ
c	113	20	0.9	398	12	BE776776	BE776776 MY-20-C-0
c	114	20	0.9	399	17	AQ074478	AQ074478 CIT-HSP-2
c	115	20	0.9	412	13	BM098910	BM098910 EBpi05_SQ
c	116	20	0.9	420	9	AJ460816	AJ460816 AJ460816
c	117	20	0.9	420	9	AJ464815	AJ464815 AJ464815
c	118	20	0.9	420	9	AJ464816	AJ464816 AJ464816
c	119	20	0.9	420	9	AJ464818	AJ464818 AJ464818
c	120	20	0.9	420	9	AJ471104	AJ471104 AJ471104
c	121	20	0.9	420	9	AJ481943	AJ481943 AJ481943
c	122	20	0.9	435	13	BM097625	BM097625 EBem04_SQ
c	123	20	0.9	438	13	BI779839	BI779839 EBem07_SQ
	124	20	0.9	438	13	BI972690	BI972690 sai81f01.
c	125	20	0.9	442	10	BB785361	BB785361 BB785361
	126	20	0.9	444	9	AA874508	AA874508 vx03b10.r
c	127	20	0.9	444	12	BF263744	BF263744 HV_CEa000
	128	20	0.9	446	13	BI498518	BI498518 sai15c01.
	129	20	0.9	448	12	BE750660	BE750660 202135 MA
	130	20	0.9	457	13	BJ073747	BJ073747 BJ073747
	131	20	0.9	460	12	BE750662	BE750662 202137 MA
	132	20	0.9	460	17	CNS06YKB	AL421137 T3 end of
c	133	20	0.9	461	14	BQ756838	BQ756838 EBem09_SQ
c	134	20	0.9	462	13	BM441169	BM441169 EBed02_SQ
c	135	20	0.9	474	13	BI777934	BI777934 EBro07_SQ
c	136	20	0.9	477	17	AQ415685	AQ415685 RPCI-11-2
c	137	20	0.9	480	9	AJ471105	AJ471105 AJ471105
c	138	20	0.9	480	9	AJ471106	AJ471106 AJ471106
c	139	20	0.9	480	9	AJ471108	AJ471108 AJ471108
c	140	20	0.9	480	9	AJ481942	AJ481942 AJ481942
c	141	20	0.9	480	9	AJ481944	AJ481944 AJ481944
c	142	20	0.9	480	9	AJ481945	AJ481945 AJ481945
c	143	20	0.9	505	13	BI947504	BI947504 HVSME1000
	144	20	0.9	514	10	AV670989	AV670989 AV670989
c	145	20	0.9	516	14	BQ768465	BQ768465 EBro08_SQ
c	146	20	0.9	521	14	BQ458499	BQ458499 HA05A22r
c	147	20	0.9	528	9	AI929849	AI929849 ul60e07.x
c	148	20	0.9	530	10	BE603260	BE603260 HVSMEh010
	149	20	0.9	537	13	BJ387354	BJ387354 BJ387354
	150	20	0.9	537	14	BQ520442	BQ520442 NISC_nl04
c	151	20	0.9	540	9	AJ471107	AJ471107 AJ471107
	152	20	0.9	546	10	AW755646	AW755646 sl06d02.y
c	153	20	0.9	548	13	BM376181	BM376181 EBma01_SQ
	154	20	0.9	549	13	BJ324340	BJ324340 BJ324340
	155	20	0.9	552	10	AW030845	AW030845 EST274135
	156	20	0.9	552	10	BE342512	BE342512 EST395356
c	157	20	0.9	559	14	BQ460503	BQ460503 HA09G03r
	158	20	0.9	561	17	AZ348237	AZ348237 1M0084K18
	159	20	0.9	566	10	AW830213	AW830213 sm24b07.y
	160	20	0.9	566	13	BJ328081	BJ328081 BJ328081
c	161	20	0.9	566	13	BM368848	BM368848 EBem08_SQ
	162	20	0.9	567	10	AW567872	AW567872 si66g08.y
c	163	20	0.9	582	13	BM369548	BM369548 EBem07_SQ
c	164	20	0.9	582	13	BM372228	BM372228 EBro03_SQ
c	165	20	0.9	582	13	BM376771	BM376771 EBem05_SQ
c	166	20	0.9	585	17	AZ906662	AZ906662 RPCI-24-2
c	167	20	0.9	587	14	BQ460364	BQ460364 HA09O16r
c	168	20	0.9	594	13	BJ448795	BJ448795 BJ448795
	169	20	0.9	595	10	BE190477	BE190477 so18a02.y
	170	20	0.9	598	10	AV922083	AV922083 AV922083
c	171	20	0.9	599	13	BM377572	BM377572 EBem04_SQ
c	172	20	0.9	608	10	AV916786	AV916786 AV916786
	173	20	0.9	610	10	BE555612	BE555612 sq04f04.y
	174	20	0.9	610	13	BJ461931	BJ461931 BJ461931
	175	20	0.9	613	13	BM527230	BM527230 sal51c11.
	176	20	0.9	616	13	BJ528838	BJ528838 BJ528838
	177	20	0.9	620	13	BJ530501	BJ530501 BJ530501
c	178	20	0.9	623	13	BI780201	BI780201 EBma08_SQ
	179	20	0.9	624	17	AQ566420	AQ566420 HS_2096_A
	180	20	0.9	629	13	BI177213	BI177213 EST518158
	181	20	0.9	632	9	AL785068	AL785068 AL785068
c	182	20	0.9	633	9	AJ433088	AJ433088 AJ433088
	183	20	0.9	644	12	BG044371	BG044371 saa27f05.
c	184	20	0.9	644	14	BQ470988	BQ470988 HX04K05r
	185	20	0.9	645	14	BQ793997	BQ793997 EST 2935
c	186	20	0.9	650	14	BQ768604	BQ768604 EBro08_SQ
c	187	20	0.9	660	17	AQ264780	AQ264780 CITBI-E1-
c	188	20	0.9	663	14	BQ768464	BQ768464 EBro08_SQ
c	189	20	0.9	671	12	BG418184	BG418184 HVSMEk002
c	190	20	0.9	672	12	BF263745	BF263745 HV_CEa000
	191	20	0.9	673	9	AI491087	AI491087 EST241796
c	192	20	0.9	676	14	BQ760071	BQ760071 EBpi07_SQ
c	193	20	0.9	677	17	AQ287739	AQ287739 nbxb0031A
	194	20	0.9	681	14	BQ510476	BQ510476 EST617891
	195	20	0.9	684	13	BI934953	BI934953 EST554842
	196	20	0.9	686	13	BM393334	BM393334 50071-2-9
	197	20	0.9	686	13	BM395422	BM395422 50072-2-9
	198	20	0.9	687	9	AU060736	AU060736 AU060736
c	199	20	0.9	687	14	BQ764827	BQ764827 EBca01_SQ
c	200	20	0.9	694	13	BJ542660	BJ542660 BJ542660
	201	20	0.9	694	14	BQ509337	BQ509337 EST616752
c	202	20	0.9	700	12	BF621942	BF621942 HVSMEa000
c	203	20	0.9	700	17	BH470168	BH470168 BOHLM19TR
c	204	20	0.9	701	12	BF622691	BF622691 HVSMEa000
	205	20	0.9	702	13	BI175951	BI175951 EST517034
	206	20	0.9	722	17	AG105109	AG105109 Pan trogl
	207	20	0.9	726	12	BF611759	BF611759 de88d02.y
c	208	20	0.9	730	13	BJ541123	BJ541123 BJ541123
	209	20	0.9	824	17	CNS04FMG	AL288529 Tetraodon
c	210	20	0.9	825	17	AZ535469	AZ535469 ENTCR48TF
	211	20	0.9	852	17	AZ669540	AZ669540 ENTLM40TR
	212	20	0.9	898	17	AZ541284	AZ541284 ENTGF32TF
c	213	20	0.9	962	17	AZ691214	AZ691214 ENTKZ88TF
	214	20	0.9	983	10	BB610861	BB610861 BB610861
	215	20	0.9	1026	17	CNS0783A	AL433484 T7 end of
	216	19	0.9	111	13	BJ387357	BJ387357 BJ387357
c	217	19	0.9	153	9	AI926243	AI926243 wo41g07.x
c	218	19	0.9	156	9	AI046086	AI046086 TENU1172
	219	19	0.9	236	9	AA926895	AA926895 om25f03.s
	220	19	0.9	237	10	BE077508	BE077508 RC1-BT060
c	221	19	0.9	244	12	BG626723	BG626723 cC-esf1cL
	222	19	0.9	244	14	N31074	N31074 yx63e03.r1
	223	19	0.9	248	10	BB046684	BB046684 BB046684
	224	19	0.9	256	10	AV361686	AV361686 AV361686
	225	19	0.9	258	17	AZ459411	AZ459411 1M0264D14

	226	19	0.9	273	17	AZ448639	AZ448639 1M0246M24
c	227	19	0.9	273	17	B31151	B31151 HS-1007-A2-
	228	19	0.9	281	10	BB080219	BB080219 BB080219
	229	19	0.9	282	9	AA372625	AA372625 EST84543
c	230	19	0.9	295	14	BQ351206	BQ351206 MR0-HT055
c	231	19	0.9	300	17	AQ003560	AQ003560 CIT-HSP-2
c	232	19	0.9	301	17	BH200283	BH200283 Sm1-49P17
c	233	19	0.9	313	10	BB280669	BB280669 BB280669
	234	19	0.9	316	9	AA601260	AA601260 no14h12.s
	235	19	0.9	323	9	AA921956	AA921956 om44g10.s
c	236	19	0.9	324	12	BG591490	BG591490 EST499332
c	237	19	0.9	331	17	BH853973	BH853973 SALK_0785
c	238	19	0.9	335	12	BG401038	BG401038 602465095
c	239	19	0.9	342	10	AV661724	AV661724 AV661724
	240	19	0.9	359	14	BQ337402	BQ337402 PM3-MT020
	241	19	0.9	360	14	C46639	C46639 C46639 Yuji
	242	19	0.9	360	14	C69716	C69716 C69716 Yuji
	243	19	0.9	360	14	D75682	D75682 CELK108F8F
c	244	19	0.9	360	17	AZ040498	AZ040498 RPCI-23-3
	245	19	0.9	367	14	BQ630662	BQ630662 sap29e12.
	246	19	0.9	370	10	AW320502	AW320502 uo19g09.y
c	247	19	0.9	382	13	BM029287	BM029287 IpSkn0035
	248	19	0.9	390	14	BQ384721	BQ384721 NISC_mn08
c	249	19	0.9	391	9	AI904340	AI904340 IL-BT050-
	250	19	0.9	395	14	BQ349893	BQ349893 MR0-HT055
	251	19	0.9	395	14	Z84125	Z84125 SSZ84125 Po
	252	19	0.9	396	17	BH805534	BH805534 1008061A0
	253	19	0.9	404	9	AA068453	AA068453 mm58c08.r
	254	19	0.9	405	14	BQ569597	BQ569597 gi133b03.
c	255	19	0.9	408	9	AI904335	AI904335 IL-BT050-
c	256	19	0.9	409	14	BQ768414	BQ768414 EBro08_SQ
	257	19	0.9	410	9	AL825240	AL825240 AL825240
	258	19	0.9	410	14	BQ666236	BQ666236 pb38c06.y
c	259	19	0.9	410	14	N46020	N46020 yy35a05.r1
	260	19	0.9	414	12	BF475449	BF475449 nac30e04.
	261	19	0.9	415	10	AW374733	AW374733 MR1-CT005
c	262	19	0.9	415	14	BQ316836	BQ316836 RC0-CT037
	263	19	0.9	415	14	BQ363330	BQ363330 RC0-ST037
	264	19	0.9	415	14	BQ363379	BQ363379 RC0-ST037
	265	19	0.9	415	14	D76943	D76943 MUS88B08 mo
c	266	19	0.9	421	12	BG876254	BG876254 RC0-CT037
c	267	19	0.9	421	14	BQ316967	BQ316967 RC0-CT037
c	268	19	0.9	426	14	BQ363331	BQ363331 RC0-ST037
c	269	19	0.9	426	14	BQ363380	BQ363380 RC0-ST037
c	270	19	0.9	437	10	AW451271	AW451271 UI-H-BI3-
	271	19	0.9	439	17	AQ776102	AQ776102 HS_5571_A
	272	19	0.9	444	9	AU183708	AU183708 AU183708
c	273	19	0.9	445	12	BF086736	BF086736 CM4-GN008
	274	19	0.9	450	10	BB861589	BB861589 BB861589
	275	19	0.9	451	17	AZ308429	AZ308429 1M0011F24
	276	19	0.9	452	10	BB858650	BB858650 BB858650
	277	19	0.9	456	9	AU181815	AU181815 AU181815
c	278	19	0.9	459	14	BQ138120	BQ138120 NF037B02P
	279	19	0.9	460	9	AU101313	AU101313 AU101313
c	280	19	0.9	462	14	N25829	N25829 yx22d01.r1
c	281	19	0.9	463	9	AL369868	AL369868 MtBA34A04
	282	19	0.9	463	13	BJ183469	BJ183469 BJ183469
	283	19	0.9	464	10	BB854093	BB854093 BB854093
	284	19	0.9	465	17	AQ494831	AQ494831 HS_5163_A
c	285	19	0.9	470	12	BF644207	BF644207 NF060C02E
c	286	19	0.9	471	12	BF250674	BF250674 pb03b07.y
c	287	19	0.9	476	10	BE075308	BE075308 PM0-BT058
c	288	19	0.9	478	17	AZ449236	AZ449236 1M0247E24
c	289	19	0.9	480	9	AJ467307	AJ467307 AJ467307
c	290	19	0.9	480	9	AJ468307	AJ468307 AJ468307
c	291	19	0.9	488	17	AZ222149	AZ222149 RPCI-23-7
c	292	19	0.9	494	14	N31401	N31401 yx67d08.r1
c	293	19	0.9	495	14	BM855120	BM855120 K-EST0137
	294	19	0.9	498	9	AL045294	AL045294 DKFZp434N
	295	19	0.9	503	13	BI336499	BI336499 AR097C04S
c	296	19	0.9	506	17	AQ187167	AQ187167 HS_3123_B
c	297	19	0.9	507	17	AQ183003	AQ183003 HS_3123_B
	298	19	0.9	508	17	AQ440700	AQ440700 HS_5114_B

c	299	19	0.9	513	12	BF651182	BF651182 NF102G02E
c	300	19	0.9	514	17	AZ741046	AZ741046 RPCI-24-1
c	301	19	0.9	518	12	BG633044	BG633044 GH18980.3
	302	19	0.9	519	17	BH363404	BH363404 CH230-154
c	303	19	0.9	520	12	BF650595	BF650595 NF098D02E
	304	19	0.9	524	9	AU161063	AU161063 AU161063
	305	19	0.9	525	12	BF252972	BF252972 EST445467
c	306	19	0.9	529	12	BG639066	BG639066 LD42989.3
	307	19	0.9	531	17	AQ516207	AQ516207 HS_5240_A
	308	19	0.9	538	12	BG062860	BG062860 L0958H08-
c	309	19	0.9	539	10	AW684103	AW684103 NF012D10N
c	310	19	0.9	540	12	BF647830	BF647830 NF012E01E
	311	19	0.9	542	12	BF089957	BF089957 MR0-MT012
	312	19	0.9	544	9	AI443739	AI443739 sa45d02.y
c	313	19	0.9	544	12	BF053868	BF053868 EST439014
	314	19	0.9	544	13	BM584363	BM584363 170006872
	315	19	0.9	547	10	AW591560	AW591560 xx14d07.x
	316	19	0.9	548	10	AW383091	AW383091 PM3-HT034
	317	19	0.9	555	10	AW500580	AW500580 UI-HF-BN0
	318	19	0.9	555	17	AZ378977	AZ378977 1M0134A04
c	319	19	0.9	557	17	TA39E12Q	AL455323 T. brucei
	320	19	0.9	560	17	AQ475011	AQ475011 CITBI-E1-
	321	19	0.9	566	13	BI796599	BI796599 H050A07 E
	322	19	0.9	573	13	BI500668	BI500668 rs80a11.y
	323	19	0.9	575	10	BB637269	BB637269 BB637269
c	324	19	0.9	576	17	TA30H07Q	AL452911 T. brucei
c	325	19	0.9	581	12	BF954450	BF954450 MR4-NN020
c	326	19	0.9	583	14	BQ837549	BQ837549 rf27g06.y
	327	19	0.9	584	17	AQ369640	AQ369640 HS_5033_B
	328	19	0.9	585	9	AL773651	AL773651 AL773651
	329	19	0.9	586	9	AA147146	AA147146 zo32c08.r
	330	19	0.9	591	9	AI900496	AI900496 sc11g05.y
	331	19	0.9	591	13	BJ108908	BJ108908 BJ108908
c	332	19	0.9	594	10	AW685720	AW685720 NF034D02N
	333	19	0.9	596	9	AA147152	AA147152 zo32d07.r
	334	19	0.9	596	13	BJ104462	BJ104462 BJ104462
	335	19	0.9	596	13	BM655643	BM655643 170006592
	336	19	0.9	598	17	AZ630181	AZ630181 1M0483C07
	337	19	0.9	601	17	AQ651115	AQ651115 Sheared D
c	338	19	0.9	604	9	AI256997	AI256997 LP04262.5
	339	19	0.9	604	9	AA239392	AA239392 mv22g09.r
	340	19	0.9	604	13	BJ066843	BJ066843 BJ066843
c	341	19	0.9	605	12	BF649570	BF649570 NF081C04E
	342	19	0.9	605	17	AZ096925	AZ096925 RPCI-23-1
	343	19	0.9	606	10	BE269182	BE269182 601184566
c	344	19	0.9	607	17	AZ735119	AZ735119 RPCI-24-7
c	345	19	0.9	611	12	BF989695	BF989695 MR2-GN012
c	346	19	0.9	613	12	BF649779	BF649779 NF081E05E
c	347	19	0.9	614	17	AZ360713	AZ360713 1M0104O03
	348	19	0.9	615	9	AL679013	AL679013 AL679013
	349	19	0.9	616	13	BJ055398	BJ055398 BJ055398
	350	19	0.9	617	13	BJ101684	BJ101684 BJ101684
	351	19	0.9	620	14	BQ792564	BQ792564 EST 8284
c	352	19	0.9	622	17	AZ292292	AZ292292 RPCI-23-1
	353	19	0.9	628	17	DR5016T	AL746096 Danio rer
	354	19	0.9	630	12	BG777332	BG777332 602664579
c	355	19	0.9	632	17	AQ949861	AQ949861 Sheared D
c	356	19	0.9	638	9	AA927343	AA927343 om69e06.s
	357	19	0.9	638	10	BB661379	BB661379 BB661379
c	358	19	0.9	639	17	AZ311045	AZ311045 1M0026I13
c	359	19	0.9	643	17	BH028378	BH028378 RPCI-24-3
	360	19	0.9	644	10	BB659068	BB659068 BB659068
c	361	19	0.9	645	9	AU235676	AU235676 AU235676
	362	19	0.9	645	17	BH615235	BH615235 BMBAC301C
c	363	19	0.9	647	10	AW940810	AW940810 GH21062.3
c	364	19	0.9	648	12	BF651260	BF651260 NF109F07E
c	365	19	0.9	649	14	BQ138261	BQ138261 NF001D08P
	366	19	0.9	652	17	AQ287800	AQ287800 nbxb0031I
c	367	19	0.9	653	12	BF646074	BF646074 NF065H11E
c	368	19	0.9	653	12	BE941825	BE941825 EST421404
c	369	19	0.9	655	10	AW685757	AW685757 NF034B09N
c	370	19	0.9	656	12	BF646628	BF646628 NF078C12E
c	371	19	0.9	657	12	BF646552	BF646552 NF076C02E

c	372	19	0.9	660	10	AW685848	AW685848 NF030H09N
	373	19	0.9	661	17	AZ600454	AZ600454 1M0418I17
	374	19	0.9	663	10	BB618815	BB618815 BB618815
c	375	19	0.9	663	12	BF645868	BF645868 NF032F08E
	376	19	0.9	663	12	BF166736	BF166736 601776251
c	377	19	0.9	665	14	BQ407620	BQ407620 GA__Ed000
	378	19	0.9	668	9	AL505041	AL505041 AL505041
	379	19	0.9	668	17	AQ475616	AQ475616 CITBI-E1-
	380	19	0.9	674	17	AZ417760	AZ417760 1M0193M06
	381	19	0.9	677	14	BM828395	BM828395 K-EST0101
c	382	19	0.9	680	12	BF645004	BF645004 NF020F10E
	383	19	0.9	680	12	BE890333	BE890333 601431788
c	384	19	0.9	680	13	BI262816	BI262816 NF091G06E
	385	19	0.9	682	17	AG042230	AG042230 Pan trogl
c	386	19	0.9	685	10	BE265000	BE265000 601193829
c	387	19	0.9	686	13	BJ143867	BJ143867 BJ143867
c	388	19	0.9	687	12	BF643529	BF643529 NF021F11E
c	389	19	0.9	690	17	AZ854117	AZ854117 2M0157A04
	390	19	0.9	691	12	BE894491	BE894491 601433091
c	391	19	0.9	693	13	BM393529	BM393529 50072-2-1
	392	19	0.9	700	9	AL503450	AL503450 AL503450
	393	19	0.9	700	9	AL504817	AL504817 AL504817
	394	19	0.9	700	12	BG592053	BG592053 EST499895
c	395	19	0.9	703	12	BG632720	BG632720 GH12767.3
c	396	19	0.9	708	17	AZ208374	AZ208374 SP_0150_A
	397	19	0.9	717	9	AA986987	AA986987 uc80h06.y
	398	19	0.9	717	14	BQ180444	BQ180444 UI-M-EX0-
	399	19	0.9	718	9	AL652289	AL652289 AL652289
	400	19	0.9	727	17	AQ051089	AQ051089 nbxb0005a
	401	19	0.9	736	17	BH543923	BH543923 BOGEF30TR
	402	19	0.9	743	9	AU123899	AU123899 AU123899
	403	19	0.9	744	13	BI102865	BI102865 602888459
c	404	19	0.9	752	12	BF046892	BF046892 EST989 Ma
c	405	19	0.9	754	13	BI550532	BI550532 603195457
	406	19	0.9	756	12	BG599810	BG599810 EST504705
	407	19	0.9	760	9	AU117406	AU117406 AU117406
	408	19	0.9	760	17	AQ935849	AQ935849 CpG2685B
	409	19	0.9	760	17	BH037225	BH037225 RPCI-24-2
c	410	19	0.9	762	12	BF302156	BF302156 602033110
	411	19	0.9	777	13	BI918073	BI918073 603181963
	412	19	0.9	779	17	CNS072W1	AL426743 clone BA0
	413	19	0.9	792	11	AK015776	AK015776 Mus muscu
c	414	19	0.9	803	12	BG645554	BG645554 EST507173
c	415	19	0.9	804	12	BG519508	BG519508 602577783
c	416	19	0.9	804	14	BQ115189	BQ115189 EST600765
	417	19	0.9	805	10	BE412836	BE412836 MCG008.B0
c	418	19	0.9	814	13	BI081767	BI081767 602879661
	419	19	0.9	816	9	AU132682	AU132682 AU132682
	420	19	0.9	821	17	AQ050907	AQ050907 nbxb0004c
	421	19	0.9	822	17	CNS035M4	AL228901 Tetraodon
c	422	19	0.9	829	13	BI824613	BI824613 603033546
c	423	19	0.9	832	12	BE865790	BE865790 601678176
c	424	19	0.9	835	17	AZ684287	AZ684287 ENTLI64TF
	425	19	0.9	837	9	AU142864	AU142864 AU142864
	426	19	0.9	843	10	BE640750	BE640750 Cri2_1_F1
	427	19	0.9	843	17	AQ576114	AQ576114 nbxb0088J
c	428	19	0.9	858	17	AZ205922	AZ205922 SP_0105_A
c	429	19	0.9	861	17	CNS06DUO	AL394294 T3_end of
c	430	19	0.9	862	14	BQ723018	BQ723018 AGENCOURT
	431	19	0.9	866	9	AU124378	AU124378 AU124378
c	432	19	0.9	873	17	CNS03OO5	AL252734 Tetraodon
c	433	19	0.9	877	17	CNS06KXW	AL403482 T7 end of
	434	19	0.9	880	12	BG255707	BG255707 602368063
c	435	19	0.9	883	17	CNS032OF	AL225096 Tetraodon
c	436	19	0.9	886	14	BQ936221	BQ936221 AGENCOURT
	437	19	0.9	890	12	BE884173	BE884173 601505928
c	438	19	0.9	895	13	BI819955	BI819955 603035314
	439	19	0.9	900	14	BQ899473	BQ899473 AGENCOURT
	440	19	0.9	906	14	BQ218557	BQ218557 AGENCOURT
	441	19	0.9	918	12	BF167886	BF167886 601775619
c	442	19	0.9	918	17	CNS039S7	AL234304 Tetraodon
	443	19	0.9	920	12	BE785343	BE785343 601478732
	444	19	0.9	923	14	BQ887781	BQ887781 AGENCOURT

c	445	19	0.9	928	12	BE898612	BE898612 601681550
	446	19	0.9	936	14	BQ680001	BQ680001 AGENCOURT
c	447	19	0.9	936	17	CNS03JOH	AL247130 Tetraodon
	448	19	0.9	949	11	AK020377	AK020377 Mus muscu
	449	19	0.9	961	14	BQ678258	BQ678258 AGENCOURT
c	450	19	0.9	962	14	BQ723669	BQ723669 AGENCOURT
c	451	19	0.9	966	12	BG484130	BG484130 602504744
	452	19	0.9	1003	13	BM471017	BM471017 AGENCOURT
	453	19	0.9	1007	17	CNS04DUN	AL286232 Tetraodon
c	454	19	0.9	1013	12	BE899378	BE899378 601681309
	455	19	0.9	1018	13	BM462593	BM462593 AGENCOURT
	456	19	0.9	1027	17	CNS03U0E	AL260519 Tetraodon
	457	19	0.9	1040	14	BM904892	BM904892 AGENCOURT
	458	19	0.9	1055	17	CNS02W3V	AL216580 Tetraodon
c	459	19	0.9	1069	17	CNS01TNP	AL166750 Tetraodon
	460	19	0.9	1082	13	BM541646	BM541646 AGENCOURT
	461	19	0.9	1082	14	BM806554	BM806554 AGENCOURT
	462	19	0.9	1089	14	BQ062618	BQ062618 AGENCOURT
	463	19	0.9	1090	13	BM550361	BM550361 AGENCOURT
	464	19	0.9	1101	17	CNS05SN7	AL352060 Tetraodon
c	465	19	0.9	1126	12	BG258475	BG258475 602379954
c	466	19	0.9	1183	13	BM546059	BM546059 AGENCOURT
c	467	19	0.9	1257	12	BG750045	BG750045 602708815
c	468	19	0.9	1452	10	BE421594	BE421594 HWM011cE.
	469	19	0.9	1671	11	BC006489	BC006489 Homo sapi
	470	19	0.9	2714	11	AF077033	AF077033 Homo sapi
	471	19	0.9	2756	11	AK010281	AK010281 Mus muscu
	472	19	0.9	4626	11	BC007668	BC007668 Homo sapi
	473	18	0.8	91	9	AI032269	AI032269 or97e05.s
c	474	18	0.8	110	10	BE319457	BE319457 NF018G10R
c	475	18	0.8	142	13	BI883146	BI883146 fn03d03.y
c	476	18	0.8	143	13	BI317943	BI317943 fq71e02.x
c	477	18	0.8	146	13	BI318187	BI318187 fq61f10.x
	478	18	0.8	150	13	BG956149	BG956149 QV1-CT075
c	479	18	0.8	163	13	BI318074	BI318074 fq58g11.x
	480	18	0.8	171	12	BE836343	BE836343 PM3-FN005
c	481	18	0.8	171	13	BI317964	BI317964 fq71g05.x
	482	18	0.8	180	9	AJ462330	AJ462330 AJ462330
c	483	18	0.8	180	10	AW126455	AW126455 614072D02
c	484	18	0.8	194	13	BI841305	BI841305 fq54a12.x
c	485	18	0.8	204	14	BQ092426	BQ092426 fz41f10.x
	486	18	0.8	207	12	BF091400	BF091400 CM0-TN003
c	487	18	0.8	207	13	BI472568	BI472568 fs01c01.x
	488	18	0.8	209	17	AZ633227	AZ633227 1M0488A19
	489	18	0.8	211	9	AV165117	AV165117 AV165117
	490	18	0.8	211	10	BE064380	BE064380 RC4-BT031
c	491	18	0.8	211	12	BG135137	BG135137 EST468029
c	492	18	0.8	212	13	BI429942	BI429942 fq77f08.x
	493	18	0.8	213	14	BQ077929	BQ077929 fy82e08.y
	494	18	0.8	215	14	BQ092923	BQ092923 fy91g06.y
c	495	18	0.8	221	13	BI707873	BI707873 fs38g12.x
	496	18	0.8	233	17	AZ304821	AZ304821 1M0005N17
	497	18	0.8	238	10	AW213648	AW213648 um58e08.y
c	498	18	0.8	239	17	AZ481633	AZ481633 1M0306E02
	499	18	0.8	240	9	AU071907	AU071907 AU071907
	500	18	0.8	240	10	AW418170	AW418170 55731 MAR
c	501	18	0.8	244	10	AV363308	AV363308 AV363308
c	502	18	0.8	248	13	BI882897	BI882897 fm99d10.x
	503	18	0.8	248	17	AQ412245	AQ412245 RPCI-11-1
	504	18	0.8	249	10	BE528509	BE528509 M79D07STM
c	505	18	0.8	250	9	AA605634	AA605634 fa31d05.s
	506	18	0.8	254	10	BB544125	BB544125 BB544125
c	507	18	0.8	254	13	BI318933	BI318933 fq64b08.x
c	508	18	0.8	254	13	BI429596	BI429596 fq49b11.x
c	509	18	0.8	255	13	BI318280	BI318280 fq67h09.x
c	510	18	0.8	255	13	BI318893	BI318893 fq62b05.x
c	511	18	0.8	255	13	BI429543	BI429543 fq48c07.x
	512	18	0.8	256	13	BM533459	BM533459 fx42b07.y
c	513	18	0.8	256	17	BH502507	BH502507 BOHPE61TF
c	514	18	0.8	258	13	BI319006	BI319006 fq66c09.x
c	515	18	0.8	261	13	BI474332	BI474332 fr95c12.x
c	516	18	0.8	263	9	AA090756	AA090756 y1678.seq
	517	18	0.8	263	13	BM155124	BM155124 fv93h06.y

	518	18	0.8	264	10	BB341316	BB341316 BB341316
	519	18	0.8	264	14	BQ330257	BQ330257 RC6-ET008
c	520	18	0.8	269	9	AL044874	AL044874 DKFZp434F
c	521	18	0.8	269	13	BI317878	BI317878 fq70b03.x
c	522	18	0.8	270	9	AA443041	AA443041 zx98g01.r
	523	18	0.8	270	17	AQ070090	AQ070090 HS_3027_B
c	524	18	0.8	271	10	AW085416	AW085416 wy64b06.x
	525	18	0.8	274	10	BB194477	BB194477 BB194477
	526	18	0.8	278	12	BF137557	BF137557 601780569
	527	18	0.8	278	13	BI884087	BI884087 fm99d10.y
	528	18	0.8	280	10	AV524608	AV524608 AV524608
	529	18	0.8	280	12	BG588558	BG588558 EST490367
	530	18	0.8	280	12	BF106648	BF106648 601823109
	531	18	0.8	282	9	AU179536	AU179536 AU179536
	532	18	0.8	282	10	AW226648	AW226648 um58g08.y
	533	18	0.8	282	10	AW715652	AW715652 g6a06nm.r
	534	18	0.8	282	10	BB180176	BB180176 BB180176
c	535	18	0.8	282	12	BG316556	BG316556 947025D05
	536	18	0.8	285	12	BF960919	BF960919 QV2-NN004
	537	18	0.8	285	14	BQ330258	BQ330258 RC6-ET008
c	538	18	0.8	286	9	AI313745	AI313745 NYZ095 Ap
	539	18	0.8	286	13	BM157398	BM157398 fw42f11.y
c	540	18	0.8	287	14	BQ092126	BQ092126 fz37e05.x
	541	18	0.8	287	14	BQ093423	BQ093423 fy97f12.y
c	542	18	0.8	289	9	AI322230	AI322230 NYZ224 Ap
	543	18	0.8	289	9	AV137163	AV137163 AV137163
	544	18	0.8	289	9	AA600577	AA600577 vm75f09.r
	545	18	0.8	290	10	BB608844	BB608844 BB608844
	546	18	0.8	292	9	AI583039	AI583039 ts03b08.x
	547	18	0.8	294	10	BB264279	BB264279 BB264279
	548	18	0.8	294	13	BM533043	BM533043 fx65d02.y
	549	18	0.8	296	10	AW758569	AW758569 NXNV_085_
	550	18	0.8	296	14	BQ132175	BQ132175 fz63g08.y
	551	18	0.8	296	14	BQ133092	BQ133092 fz37e05.y
c	552	18	0.8	298	10	AW393645	AW393645 RC5-ST031
c	553	18	0.8	300	17	TA97E08P	AL461035 T. brucei
c	554	18	0.8	301	9	AI313720	AI313720 NYZ067 Ap
c	555	18	0.8	301	10	BB477818	BB477818 BB477818
c	556	18	0.8	303	9	AI330411	AI330411 fa92d04.x
	557	18	0.8	303	13	BM157365	BM157365 fw41f12.y
	558	18	0.8	304	9	AA237459	AA237459 mw95h02.r
c	559	18	0.8	304	13	BI840614	BI840614 fq52f08.x
	560	18	0.8	305	10	AW965821	AW965821 EST377894
	561	18	0.8	309	9	AA736716	AA736716 nw53f06.s
	562	18	0.8	310	9	AI583645	AI583645 tt74d07.x
c	563	18	0.8	311	9	AI444415	AI444415 fb38b09.x
c	564	18	0.8	311	9	AU260030	AU260030 AU260030
	565	18	0.8	311	13	BM159070	BM159070 fw25d03.y
	566	18	0.8	313	9	AL384660	AL384660 MtBC23F04
	567	18	0.8	314	9	AA382146	AA382146 EST95418
c	568	18	0.8	314	12	BF756577	BF756577 QV3-CT055
c	569	18	0.8	314	13	BI429538	BI429538 fq48b08.x
c	570	18	0.8	315	10	BB501934	BB501934 BB501934
c	571	18	0.8	315	13	BI841468	BI841468 fq56e11.x
	572	18	0.8	317	9	AI437135	AI437135 fb38b09.y
c	573	18	0.8	318	10	BB285877	BB285877 BB285877
c	574	18	0.8	319	12	BG892640	BG892640 fq77e05.x
c	575	18	0.8	321	9	AI442211	AI442211 sa49b05.y
c	576	18	0.8	321	13	BI429642	BI429642 fq50b03.x
	577	18	0.8	322	9	AI277490	AI277490 ql96c11.x
	578	18	0.8	322	10	BB530214	BB530214 BB530214
	579	18	0.8	323	9	AI561460	AI561460 vj94a12.x
	580	18	0.8	326	10	AW321497	AW321497 uo32g06.y
c	581	18	0.8	327	10	AW383180	AW383180 PM3-HT034
c	582	18	0.8	328	9	AI239947	AI239947 qh44a06.x
	583	18	0.8	332	12	BF569617	BF569617 602186130
c	584	18	0.8	333	9	AI764604	AI764604 UI-R-Y0-a
c	585	18	0.8	334	9	AU023282	AU023282 AU023282
	586	18	0.8	334	10	AW447543	AW447543 89305 MAR
	587	18	0.8	336	10	BE521449	BE521449 M19F5STM
c	588	18	0.8	337	12	BF417301	BF417301 UI-R-CN0-
c	589	18	0.8	337	13	BI707870	BI707870 fs38g08.x
c	590	18	0.8	338	9	AA441801	AA441801 zw62d02.s

c	591	18	0.8	340	9	AI396637	AI396637 fb15d04.x
	592	18	0.8	340	17	AQ646105	AQ646105 RPCI93-Ec
	593	18	0.8	341	17	AQ008733	AQ008733 RPCI11-22
c	594	18	0.8	342	12	BG303148	BG303148 f161f03.x
	595	18	0.8	342	17	AQ848303	AQ848303 LMAJFV1_1
	596	18	0.8	343	17	AZ826972	AZ826972 2M0103P04
	597	18	0.8	344	10	BB852967	BB852967 BB852967
c	598	18	0.8	346	9	AI332078	AI332078 fa93c01.x
c	599	18	0.8	346	9	AJ282285	AJ282285 4A3A-AAS-
	600	18	0.8	347	12	BF511970	BF511970 UI-H-BI4-
c	601	18	0.8	348	9	AI313714	AI313714 NYZ061 Ap
c	602	18	0.8	349	13	BI318114	BI318114 fq59f07.x
c	603	18	0.8	349	13	BI472624	BI472624 fs02c02.x
	604	18	0.8	349	14	BQ801082	BQ801082 WHE2810_B
c	605	18	0.8	350	13	BI708111	BI708111 fs42b06.x
c	606	18	0.8	350	14	BQ348857	BQ348857 CM4-HT039
c	607	18	0.8	352	14	D74652	D74652 CELK083F9F
c	608	18	0.8	354	10	AW778714	AW778714 fk40g10.x
c	609	18	0.8	355	10	AW116676	AW116676 fi18e09.x
c	610	18	0.8	355	13	BI982437	BI982437 ft99b06.x
c	611	18	0.8	357	13	BI846434	BI846434 fq47c02.x
c	612	18	0.8	357	13	BM081981	BM081981 fu97g02.x
	613	18	0.8	357	17	AQ007140	AQ007140 CIT-HSP-2
c	614	18	0.8	358	14	BQ092335	BQ092335 fz40d07.x
c	615	18	0.8	358	17	AQ044991	AQ044991 RPCI11-33
c	616	18	0.8	359	17	AZ301682	AZ301682 Gm_UMb001
	617	18	0.8	360	9	AI396801	AI396801 fb14c06.y
c	618	18	0.8	360	13	BI318156	BI318156 fq60h01.x
	619	18	0.8	360	17	AQ956947	AQ956947 LERAM94TR
	620	18	0.8	361	9	AJ280664	AJ280664 4A3A-AAS-
c	621	18	0.8	362	9	AI626345	AI626345 fc13h08.x
	622	18	0.8	362	13	BI882567	BI882567 fn03d03.x
c	623	18	0.8	363	13	BI317901	BI317901 fq70f01.x
c	624	18	0.8	364	17	CNS03W1N	AL263156 Tetraodon
c	625	18	0.8	365	10	AW567052	AW567052 fk32c07.x
c	626	18	0.8	365	10	AW567473	AW567473 fk38h11.x
c	627	18	0.8	365	10	AW567544	AW567544 fk29a05.x
c	628	18	0.8	366	10	AW510124	AW510124 fk16h07.x
c	629	18	0.8	366	10	AW567005	AW567005 fk30f08.x
c	630	18	0.8	368	10	AW566953	AW566953 fk19d07.x
	631	18	0.8	368	12	BF514196	BF514196 UI-H-BW1-
c	632	18	0.8	368	14	BM778009	BM778009 fy26e06.x
c	633	18	0.8	369	10	AV525109	AV525109 AV525109
	634	18	0.8	369	17	AQ127029	AQ127029 HS_3041_B
c	635	18	0.8	370	10	AW203433	AW203433 sf30b11.y
c	636	18	0.8	371	13	BI866174	BI866174 fu01c07.x
c	637	18	0.8	372	9	AA517969	AA517969 vi01h03.r
c	638	18	0.8	372	10	AW567002	AW567002 fk30f05.x
c	639	18	0.8	372	14	W66656	W66656 me23d12.r1
c	640	18	0.8	375	10	AW154230	AW154230 fi24b11.x
c	641	18	0.8	375	10	AW154555	AW154555 fi28e05.x
c	642	18	0.8	375	10	AW343602	AW343602 fi46d06.x
c	643	18	0.8	375	12	BG307545	BG307545 f159f02.x
	644	18	0.8	375	12	BE751713	BE751713 203917 MA
	645	18	0.8	375	12	BE752026	BE752026 204301 MA
c	646	18	0.8	376	9	AI292380	AI292380 NYZ024 Ap
c	647	18	0.8	377	12	BG303495	BG303495 f156e03.x
c	648	18	0.8	378	9	AA317819	AA317819 EST19877
	649	18	0.8	378	14	BM862424	BM862424 mgcm003xG
c	650	18	0.8	379	9	AI376394	AI376394 tc29e08.x
c	651	18	0.8	379	12	BF779381	BF779381 3191-90 h
c	652	18	0.8	379	13	BI844045	BI844045 fs32g12.x
	653	18	0.8	380	9	AI353417	AI353417 zeh0471.s
c	654	18	0.8	380	13	BI877638	BI877638 f162g07.x
	655	18	0.8	381	10	AW485383	AW485383 65290 MAR
c	656	18	0.8	382	13	BI317890	BI317890 fq70d05.x
c	657	18	0.8	384	9	AA605626	AA605626 fa30h10.s
	658	18	0.8	384	12	BG090853	BG090853 mac11h12.
c	659	18	0.8	386	10	AW154494	AW154494 fi27g07.x
c	660	18	0.8	386	13	BI878604	BI878604 f180d12.x
c	661	18	0.8	387	9	AI332081	AI332081 fa93c06.x
c	662	18	0.8	387	17	BH741678	BH741678 gt36b02.g
c	663	18	0.8	388	9	AI322208	AI322208 NYZ168 Ap

c	664	18	0.8	388	13	BI474394	BI474394 fr96d05.x
	665	18	0.8	388	17	BH018315	BH018315 L142k.d_H
c	666	18	0.8	390	10	AV591715	AV591715 AV591715
	667	18	0.8	390	10	AW886192	AW886192 RC5-OT007
c	668	18	0.8	391	9	AI461392	AI461392 fb44h03.x
	669	18	0.8	392	9	AJ503218	AJ503218 AJ503218
	670	18	0.8	392	12	BG798218	BG798218 ic07b01.y
	671	18	0.8	392	14	R08011	R08011 yf17a07.s1
c	672	18	0.8	393	9	AI330884	AI330884 fb05c11.x
c	673	18	0.8	393	9	AI331376	AI331376 fa98a11.x
c	674	18	0.8	393	9	AI385060	AI385060 fb06f05.x
	675	18	0.8	393	9	AA441862	AA441862 zw62d02.r
c	676	18	0.8	393	10	AW175461	AW175461 fi25d09.x
c	677	18	0.8	393	12	BG882969	BG882969 fp32b06.x
	678	18	0.8	393	14	BM797887	BM797887 K-EST0081
c	679	18	0.8	393	14	BQ093716	BQ093716 fz50h08.x
	680	18	0.8	393	17	AZ304852	AZ304852 1M0005F19
	681	18	0.8	394	9	AA954411	AA954411 on85a08.s
c	682	18	0.8	394	9	AI332065	AI332065 fa93a11.x
	683	18	0.8	394	9	AI882689	AI882689 fb15b10.y
c	684	18	0.8	394	10	AW420437	AW420437 fj81e12.x
c	685	18	0.8	394	10	AW567462	AW567462 fk38g07.x
c	686	18	0.8	394	14	BQ616909	BQ616909 fab01f12.
c	687	18	0.8	395	9	AI396621	AI396621 fb15b10.x
	688	18	0.8	395	10	BE524893	BE524893 M55F3STM
	689	18	0.8	395	17	BH071136	BH071136 RPCI-24-2
c	690	18	0.8	395	17	AQ409333	AQ409333 HS_5067_A
c	691	18	0.8	396	9	AI331263	AI331263 fa97b05.x
	692	18	0.8	396	10	BE597308	BE597308 PI1_71_H0
	693	18	0.8	396	17	AQ009088	AQ009088 RPCI11-22
	694	18	0.8	397	12	BE753153	BE753153 206068 MA
	695	18	0.8	397	13	BI309820	BI309820 EST531157
c	696	18	0.8	397	13	BI841412	BI841412 fq55f10.x
c	697	18	0.8	397	13	BI843467	BI843467 ft90a06.x
c	698	18	0.8	398	9	AI330336	AI330336 fa91a11.x
c	699	18	0.8	398	9	AI330869	AI330869 fb05b05.x
c	700	18	0.8	398	10	AW778715	AW778715 fk40g11.x
	701	18	0.8	398	10	AW886196	AW886196 RC5-OT007
c	702	18	0.8	399	9	AI385082	AI385082 fb06h09.x
c	703	18	0.8	400	9	AI330522	AI330522 fa92g07.x
c	704	18	0.8	401	9	AI331795	AI331795 fb01e04.x
	705	18	0.8	401	9	AA101998	AA101998 zl78a06.r
c	706	18	0.8	402	10	AW778710	AW778710 fk40g05.x
c	707	18	0.8	404	13	BM081889	BM081889 fu96c02.x
c	708	18	0.8	404	17	AZ657923	AZ657923 1M0534K24
c	709	18	0.8	405	10	AW744362	AW744362 ur31d09.y
c	710	18	0.8	405	13	BI429892	BI429892 fq76e11.x
c	711	18	0.8	405	14	H86709	H86709 yt08a10.r1
	712	18	0.8	406	12	BF559749	BF559749 UI-R-E1-g
c	713	18	0.8	406	13	BI429890	BI429890 fq76e09.x
	714	18	0.8	407	9	AU174950	AU174950 AU174950
c	715	18	0.8	408	17	AQ787388	AQ787388 HS_3224_B
c	716	18	0.8	409	9	AI415823	AI415823 fb34e05.x
c	717	18	0.8	411	12	BE991893	BE991893 UI-M-BZ1-
c	718	18	0.8	411	13	BM289291	BM289291 530866 MA
	719	18	0.8	412	14	BM862423	BM862423 mgcm003xG
c	720	18	0.8	414	14	C73753	C73753 C73753 Rice
c	721	18	0.8	415	14	BQ451845	BQ451845 PfESToab0
c	722	18	0.8	417	12	BG892532	BG892532 fq74f07.x
c	723	18	0.8	417	12	BF226166	BF226166 uz48d08.x
c	724	18	0.8	417	13	BI275463	BI275463 UI-R-CX0-
	725	18	0.8	418	10	BE366814	BE366814 PI1_41_G0
c	726	18	0.8	418	12	BG892651	BG892651 fq77f08.x
c	727	18	0.8	418	13	BI334245	BI334245 602997754
c	728	18	0.8	419	10	AW710611	AW710611 e5a01ne.f
c	729	18	0.8	419	13	BI318914	BI318914 fq63b04.x
c	730	18	0.8	419	13	BI318958	BI318958 fq65a06.x
c	731	18	0.8	419	13	BJ479399	BJ479399 BJ479399
	732	18	0.8	422	9	AI288259	AI288259 ql80c09.x
c	733	18	0.8	422	9	AI397482	AI397482 fb18d09.x
c	734	18	0.8	422	10	AW614916	AW614916 hh70f01.x
c	735	18	0.8	422	14	R82176	R82176 yj17a01.s1
c	736	18	0.8	423	14	BM671510	BM671510 UI-E-CK1-

	737	18	0.8	425	17	B86909	B86909 RPCI11-26J1
c	738	18	0.8	426	13	BI317924	BI317924 fq71c02.x
c	739	18	0.8	426	13	BI882960	BI882960 fn01d06.x
	740	18	0.8	428	10	AW604922	AW604922 RC0-CT037
	741	18	0.8	428	10	AW862374	AW862374 RC0-CT037
c	742	18	0.8	428	14	BQ515564	BQ515564 EST622979
	743	18	0.8	429	13	BI433161	BI433161 EST535922
	744	18	0.8	429	14	BQ515563	BQ515563 EST622978
c	745	18	0.8	429	17	BH223514	BH223514 1006113D0
	746	18	0.8	429	17	FR0044532	AL132024 Fugu rubr
	747	18	0.8	430	9	AI741515	AI741515 wg21c01.x
	748	18	0.8	430	9	AI943751	AI943751 618045A12
	749	18	0.8	430	13	BM028966	BM028966 IpSkn0182
c	750	18	0.8	432	9	AA717831	AA717831 vt98h03.r
c	751	18	0.8	432	13	BI562946	BI562946 fq67h09.x
c	752	18	0.8	432	13	BI864260	BI864260 ft93g01.x
	753	18	0.8	432	17	AQ728217	AQ728217 HS_5457_B
c	754	18	0.8	433	13	BI841415	BI841415 fq55g02.x
	755	18	0.8	433	14	BM856200	BM856200 K-EST0139
c	756	18	0.8	433	14	BQ393795	BQ393795 NISC_ng05
	757	18	0.8	434	14	R82218	R82218 yj17a01.r1
c	758	18	0.8	437	12	BF823837	BF823837 NcEST3a27
	759	18	0.8	438	10	AW677615	AW677615 DG1_8_B10
c	760	18	0.8	438	13	BI864373	BI864373 ft95b06.x
	761	18	0.8	438	14	BQ167623	BQ167623 WHE0081_C
c	762	18	0.8	439	10	AW525853	AW525853 UI-R-BJ0p
	763	18	0.8	441	10	AW677571	AW677571 DG1_8_H09
c	764	18	0.8	441	17	AZ049883	AZ049883 GSSTc1072
	765	18	0.8	442	10	BE600291	BE600291 PI1_94_F0
	766	18	0.8	442	13	BI294608	BI294608 UI-R-DK0-
c	767	18	0.8	442	17	B22546	B22546 F20G5TF IGF
	768	18	0.8	443	10	AW975360	AW975360 EST387468
c	769	18	0.8	443	13	BM386075	BM386075 UI-R-DN1-
	770	18	0.8	444	9	AA545284	AA545284 vj94a12.s
	771	18	0.8	444	14	W57704	W57704 zd18b02.s1
	772	18	0.8	445	10	AW886190	AW886190 RC5-OT007
c	773	18	0.8	445	13	BI707672	BI707672 fs36b08.x
	774	18	0.8	445	14	W98555	W98555 mg14d02.r1
c	775	18	0.8	445	17	AZ352576	AZ352576 1M0091A09
c	776	18	0.8	446	9	AI506751	AI506751 vm56e05.x
c	777	18	0.8	447	9	AI752740	AI752740 cn18e02.x
	778	18	0.8	447	12	BE972319	BE972319 601652305
c	779	18	0.8	447	13	BM081796	BM081796 fu94h01.x
	780	18	0.8	447	17	B84170	B84170 RPCI11-22G7
c	781	18	0.8	448	9	AA606120	AA606120 fa16a07.s
c	782	18	0.8	448	13	BI702078	BI702078 sag41g09.
c	783	18	0.8	448	17	AQ504349	AQ504349 RPCI-11-2
	784	18	0.8	449	10	AW669419	AW669419 112783 MA
c	785	18	0.8	449	14	BQ576977	BQ576977 PfESToab1
c	786	18	0.8	451	13	BI474459	BI474459 fr98d09.x
	787	18	0.8	452	10	BE363310	BE363310 WS1_61_A0
c	788	18	0.8	452	13	BI864477	BI864477 ft96e08.x
c	789	18	0.8	453	13	BI841550	BI841550 fq57e11.x
c	790	18	0.8	453	13	BI843585	BI843585 ft91f01.x
c	791	18	0.8	453	13	BI864445	BI864445 ft96a12.x
	792	18	0.8	453	17	AQ816232	AQ816232 HS_5443_B
c	793	18	0.8	454	9	AA605954	AA605954 fa14b04.s
c	794	18	0.8	454	10	AW940655	AW940655 GH17344.3
	795	18	0.8	454	10	BB861395	BB861395 BB861395
c	796	18	0.8	454	13	BI429930	BI429930 fq77c12.x
c	797	18	0.8	454	13	BI474409	BI474409 fr96f09.x
c	798	18	0.8	454	13	BI707719	BI707719 fs36f03.x
c	799	18	0.8	454	13	BI707985	BI707985 fs40d06.x
	800	18	0.8	454	13	BI784634	BI784634 saf57d06.
c	801	18	0.8	454	13	BI844015	BI844015 fs32d10.x
c	802	18	0.8	454	13	BI982460	BI982460 ft99e01.x
	803	18	0.8	454	14	T97908	T97908 ye54c11.s1
c	804	18	0.8	455	9	AI444488	AI444488 fb39h05.x
c	805	18	0.8	455	13	BI708187	BI708187 fs43c09.x
c	806	18	0.8	455	13	BI843989	BI843989 fs32a08.x
c	807	18	0.8	455	13	BI864500	BI864500 ft96g11.x
c	808	18	0.8	455	13	BI866074	BI866074 ft98a03.x
c	809	18	0.8	455	13	BJ190987	BJ190987 BJ190987

c	810	18	0.8	456	9	AI635141	AI635141 tz02e01.x
c	811	18	0.8	456	9	AI953666	AI953666 wq23h10.x
c	812	18	0.8	456	9	AA606125	AA606125 fa16a12.s
	813	18	0.8	456	10	AW027653	AW027653 wv60a08.x
c	814	18	0.8	456	10	AW993958	AW993958 RC1-BN003
c	815	18	0.8	456	13	BI472616	BI472616 fs02a07.x
c	816	18	0.8	456	17	BH779865	BH779865 fzmb014f0
	817	18	0.8	457	9	AI034007	AI034007 ow13c08.x
c	818	18	0.8	457	9	AI545279	AI545279 fb80b04.x
c	819	18	0.8	457	9	AA605966	AA605966 fa14c04.s
c	820	18	0.8	457	10	BE017613	BE017613 fk81a06.x
c	821	18	0.8	457	13	BI472718	BI472718 fs03h02.x
c	822	18	0.8	457	13	BJ190757	BJ190757 BJ190757
c	823	18	0.8	458	10	AW782678	AW782678 db22a06.x
c	824	18	0.8	458	10	BE200779	BE200779 fk91b05.x
	825	18	0.8	458	12	BG872740	BG872740 602793834
c	826	18	0.8	458	13	BI707808	BI707808 fs37g06.x
c	827	18	0.8	458	13	BI841572	BI841572 fq57h04.x
c	828	18	0.8	459	9	AA683504	AA683504 zl55d11.s
	829	18	0.8	459	10	BE652927	BE652927 UI-M-AM1-
c	830	18	0.8	460	10	BE686309	BE686309 uv74d06.x
c	831	18	0.8	461	13	BI864239	BI864239 ft93d11.x
	832	18	0.8	461	13	BM108642	BM108642 EST556178
c	833	18	0.8	461	17	AQ150788	AQ150788 HS_3201_B
c	834	18	0.8	462	9	AA606148	AA606148 fa16d02.s
c	835	18	0.8	462	13	BI707959	BI707959 fs40a10.x
c	836	18	0.8	462	13	BI864508	BI864508 ft96h09.x
	837	18	0.8	462	13	BM086133	BM086133 sah35h07.
c	838	18	0.8	463	10	AW360165	AW360165 47606 MAR
c	839	18	0.8	463	13	BI846377	BI846377 fq46d04.x
c	840	18	0.8	463	17	AQ613183	AQ613183 HS_5133_A
c	841	18	0.8	464	9	AI444208	AI444208 fb24f07.x
c	842	18	0.8	464	9	AU101502	AU101502 AU101502
	843	18	0.8	464	10	AW188295	AW188295 xj94h01.x
c	844	18	0.8	464	13	BI318105	BI318105 fq59d10.x
c	845	18	0.8	464	13	BI474447	BI474447 fr98c04.x
c	846	18	0.8	464	13	BI864310	BI864310 ft94d02.x
c	847	18	0.8	465	13	BI302308	BI302308 UI-R-DM0-
c	848	18	0.8	467	10	AW708906	AW708906 dld12ne.f
c	849	18	0.8	467	13	BI843438	BI843438 ft89f02.x
c	850	18	0.8	467	14	BM693218	BM693218 UI-E-CK1-
	851	18	0.8	468	9	AA498581	AA498581 vh44g03.r
	852	18	0.8	469	9	AI957424	AI957424 fc90h09.x
c	853	18	0.8	469	14	BQ092432	BQ092432 fz41g04.x
c	854	18	0.8	470	9	AI470323	AI470323 tj42b09.x
c	855	18	0.8	470	13	BI841381	BI841381 fq55b07.x
c	856	18	0.8	471	13	BJ177334	BJ177334 BJ177334
	857	18	0.8	471	17	AQ155085	AQ155085 HS_3039_B
c	858	18	0.8	472	12	BG892627	BG892627 fq77c12.x
c	859	18	0.8	472	13	BI846432	BI846432 fq47b11.x
c	860	18	0.8	473	12	BE945537	BE945537 UI-M-BZ0-
c	861	18	0.8	473	13	BI317992	BI317992 fq73c06.x
c	862	18	0.8	473	13	BI472710	BI472710 fs03f09.x
	863	18	0.8	473	14	N27957	N27957 yy03f02.s1
c	864	18	0.8	474	12	BF067079	BF067079 st48c03.y
	865	18	0.8	474	14	BQ301753	BQ301753 QV0-BT007
c	866	18	0.8	475	13	BI474412	BI474412 fr96g01.x
c	867	18	0.8	475	13	BI786002	BI786002 sai31f06.
c	868	18	0.8	475	13	BI866039	BI866039 ft97e04.x
c	869	18	0.8	476	13	BI318950	BI318950 fq64f09.x
c	870	18	0.8	476	17	AQ736598	AQ736598 HS_3212_A
c	871	18	0.8	476	17	AQ229010	AQ229010 HS_2019_A
	872	18	0.8	477	12	BG839624	BG839624 Gm01_04g1
c	873	18	0.8	477	17	BH690258	BH690258 BOMIS36TF
c	874	18	0.8	478	9	AI037566	AI037566 ub52h06.r
	875	18	0.8	478	9	AL369832	AL369832 MtBA33G04
	876	18	0.8	478	14	BQ133448	BQ133448 fz41g04.y
	877	18	0.8	479	10	AW921916	AW921916 EST353220
c	878	18	0.8	479	10	BE119706	BE119706 UI-R-CA0-
	879	18	0.8	481	9	AA277000	AA277000 vc46b11.r
c	880	18	0.8	482	10	BE609878	BE609878 sq46c09.y
c	881	18	0.8	482	13	BI562978	BI562978 fq68d08.x
c	882	18	0.8	483	13	BI474895	BI474895 fp54b10.x
c	883	18	0.8	484	17	AQ209144	AQ209144 HS_3227_A
c	884	18	0.8	485	10	AW114083	AW114083 rs68d03.y
c	885	18	0.8	486	10	BE201639	BE201639 fk96c05.x
c	886	18	0.8	488	9	AA606196	AA606196 fa16f11.s
	887	18	0.8	488	14	N27823	N27823 yy02a02.s1
c	888	18	0.8	489	13	BI707945	BI707945 fs39h03.x
c	889	18	0.8	489	13	BI841406	BI841406 fq55e10.x
c	890	18	0.8	489	13	BI844283	BI844283 fs35h09.x
c	891	18	0.8	489	13	BI864244	BI864244 ft93e04.x
c	892	18	0.8	489	13	BI882231	BI882231 fm90e08.x
c	893	18	0.8	489	13	BI982433	BI982433 ft99b02.x
c	894	18	0.8	489	13	BM081179	BM081179 fu90d11.x
c	895	18	0.8	489	13	BM081977	BM081977 fu97f06.x
c	896	18	0.8	489	13	BM082000	BM082000 fu97g12.x
	897	18	0.8	489	14	N94866	N94866 yy63b01.r1
	898	18	0.8	490	9	AL388250	AL388250 MtBC47E04
c	899	18	0.8	490	12	BG892600	BG892600 fq75h03.x
c	900	18	0.8	490	13	BI317937	BI317937 fq71d05.x
c	901	18	0.8	490	13	BI318104	BI318104 fq59d09.x
c	902	18	0.8	490	13	BI318903	BI318903 fq62e11.x
	903	18	0.8	490	17	AQ515978	AQ515978 HS_5240_A
c	904	18	0.8	491	10	BE509387	BE509387 dc10b01.y
	905	18	0.8	491	12	BG134223	BG134223 EST467115
c	906	18	0.8	491	13	BI429633	BI429633 fq49h12.x
c	907	18	0.8	492	9	AI202019	AI202019 qi43a08.x
	908	18	0.8	492	13	BI893429	BI893429 sai66a07.
c	909	18	0.8	494	10	AW377889	AW377889 MR0-HT020
c	910	18	0.8	494	12	BG649969	BG649969 sad90a12.
c	911	18	0.8	494	14	BQ338174	BQ338174 MR0-NN014
c	912	18	0.8	495	10	BE575950	BE575950 dc65h12.x
c	913	18	0.8	495	13	BI273881	BI273881 UI-R-CW0-
c	914	18	0.8	495	13	BJ250774	BJ250774 BJ250774
	915	18	0.8	495	14	BM779985	BM779985 EST590561
c	916	18	0.8	496	10	AW567167	AW567167 fk22a12.x
	917	18	0.8	496	12	BF706111	BF706111 280355 MA
c	918	18	0.8	496	13	BI429782	BI429782 fq74f07.x
	919	18	0.8	496	17	BH018650	BH018650 L189e.d_H
c	920	18	0.8	497	17	AQ396796	AQ396796 mgxb0015N
	921	18	0.8	497	17	AQ472411	AQ472411 CITBI-E1-
	922	18	0.8	498	9	AA498470	AA498470 vh40d06.r
	923	18	0.8	498	12	BF472933	BF472933 UI-M-BZ0-
	924	18	0.8	498	12	BF431198	BF431198 7o09b09.x
c	925	18	0.8	498	13	BI708203	BI708203 fs43e06.x
c	926	18	0.8	498	13	BM081906	BM081906 fu96e01.x
c	927	18	0.8	498	17	AZ462809	AZ462809 1M0271L18
	928	18	0.8	499	13	BI320679	BI320679 sae47c01.
c	929	18	0.8	499	13	BI707898	BI707898 fs39b10.x
c	930	18	0.8	500	13	BI866137	BI866137 ft98g10.x
	931	18	0.8	500	14	BQ704032	BQ704032 946109F03
	932	18	0.8	501	13	BI315675	BI315675 saf75g05.
c	933	18	0.8	501	13	BI429606	BI429606 fq49d10.x
c	934	18	0.8	501	13	BI472706	BI472706 fs03f05.x
c	935	18	0.8	502	10	BE025662	BE025662 db26e06.y
c	936	18	0.8	502	10	BE605732	BE605732 fl17h07.x
c	937	18	0.8	502	11	AF445320	AF445320 Penaeus m
	938	18	0.8	504	9	AI745254	AI745254 wg10g08.x
c	939	18	0.8	504	13	BI699407	BI699407 sag36f06.
c	940	18	0.8	504	13	BI707995	BI707995 fs40e08.x
c	941	18	0.8	505	9	AI765222	AI765222 wi72g05.x
c	942	18	0.8	505	12	BG726230	BG726230 sae13b02.
c	943	18	0.8	505	13	BM494075	BM494075 IpCGBr1_1
c	944	18	0.8	505	14	BQ092229	BQ092229 fz38h11.x
c	945	18	0.8	505	14	BQ131577	BQ131577 fz51a09.x
c	946	18	0.8	505	14	BQ131768	BQ131768 fz54a08.x
c	947	18	0.8	506	9	AI477309	AI477309 fb51g12.x
c	948	18	0.8	506	9	AI807690	AI807690 wf49h04.x
c	949	18	0.8	506	10	BB705224	BB705224 BB705224
c	950	18	0.8	506	10	BE201571	BE201571 fk95c01.x
	951	18	0.8	506	10	BE319749	BE319749 NF020F04R
c	952	18	0.8	506	10	BE605636	BE605636 fl14e09.x
	953	18	0.8	506	12	BG364372	BG364372 dab87a06.
c	954	18	0.8	506	12	BG726218	BG726218 sae12h11.
	955	18	0.8	506	13	BJ204367	BJ204367 BJ204367

```
c 956   18   0.8   507  10  AW460149        AW460149 si10h03.y
c 957   18   0.8   507  10  BE016311        BE016311 fk71e07.x
c 958   18   0.8   507  10  BE016445        BE016445 fk73b06.x
c 959   18   0.8   507  10  BE202220        BE202220 fl08d09.x
c 960   18   0.8   508   9  AI793645        AI793645 fc50a04.x
  961   18   0.8   508  12  BF636812        BF636812 NF072G11L
c 962   18   0.8   509  10  AW232751        AW232751 fj29h02.x
c 963   18   0.8   509  10  AW777893        AW777893 fk52d04.x
c 964   18   0.8   509  13  BI841521        BI841521 fq57c02.x
c 965   18   0.8   509  13  BI843489        BI843489 ft90c10.x
c 966   18   0.8   509  13  BI844065        BI844065 fs33a12.x
c 967   18   0.8   509  13  BI864358        BI864358 ft94h09.x
c 968   18   0.8   509  13  BM081353        BM081353 fu92g08.x
  969   18   0.8   509  17  AZ447938        AZ447938 1M0245P14
  970   18   0.8   510   9  AI325900        AI325900 mg14d02.y
c 971   18   0.8   510   9  AA142194        AA142194 CK00327.3
c 972   18   0.8   510  10  AV841870        AV841870 AV841870
c 973   18   0.8   510  13  BI318100        BI318100 fq59c12.x
c 974   18   0.8   510  13  BI707686        BI707686 fs36b03.x
c 975   18   0.8   510  13  BI707760        BI707760 fs37b05.x
c 976   18   0.8   510  13  BI864432        BI864432 ft95h07.x
c 977   18   0.8   511  10  BE046204        BE046204 hq91d06.x
c 978   18   0.8   512  13  BM081772        BM081772 fu94e02.x
  979   18   0.8   512  14  H51966          H51966 yq32b04.r1
  980   18   0.8   512  17  AZ613011        AZ613011 1M0441K10
c 981   18   0.8   513  13  BI840600        BI840600 fq52d11.x
c 982   18   0.8   513  13  BI841322        BI841322 fq54c09.x
c 983   18   0.8   513  13  BI841420        BI841420 fq55g09.x
c 984   18   0.8   513  13  BI841449        BI841449 fq56b07.x
c 985   18   0.8   513  13  BI843415        BI843415 ft89c10.x
c 986   18   0.8   513  13  BI843574        BI843574 ft91d05.x
c 987   18   0.8   513  13  BI844266        BI844266 fs35g11.x
c 988   18   0.8   513  13  BI864303        BI864303 ft94c07.x
c 989   18   0.8   513  13  BI864305        BI864305 ft94c09.x
c 990   18   0.8   513  13  BI864370        BI864370 ft95b03.x
c 991   18   0.8   513  13  BI881873        BI881873 fm83h03.x
c 992   18   0.8   514   9  AI545280        AI545280 fb80b06.x
  993   18   0.8   514  10  AV524379        AV524379 AV524379
c 994   18   0.8   514  10  BE597261        BE597261 PI1_70_B0
c 995   18   0.8   514  13  BI318027        BI318027 fq58a07.x
c 996   18   0.8   514  13  BM081287        BM081287 fu91h07.x
  997   18   0.8   515  13  BI097883        BI097883 IP1_29_C0
c 998   18   0.8   515  13  BI883036        BI883036 fn02f06.x
  999   18   0.8   516   9  AA238108        AA238108 mw97a08.r
c1000   18   0.8   517   9  AI667310        AI667310 fc39a12.x
```

ALIGNMENTS

```
RESULT 1
BC030638
LOCUS       BC030638               2523 bp    mRNA    linear   HTC 21-MAY-2002
DEFINITION  Homo sapiens, VASA protein, clone IMAGE:4822528, mRNA.
ACCESSION   BC030638
VERSION     BC030638.1  GI:21040391
KEYWORDS    HTC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2523)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 33 Row: d Column: 8
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 9507236
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..2523
                     /organism="Homo sapiens"
                     /db_xref="LocusID:54514"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4822528"
                     /tissue_type="Testis"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
BASE COUNT      780 a    437 c    616 g    690 t
ORIGIN

  Query Match             83.1%;  Score 1806;  DB 11;  Length 2523;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2126;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 163

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 223

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 283

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 343

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 403

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 463

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 523

Qy        427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 583

Qy        487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 643

Qy        547 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 703

Qy        607 GGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCT 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        704 GGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCT 763

Qy        667 GGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCA 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 GGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCA 823

Qy        727 AAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTAT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 AAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTAT 883

Qy        787 CAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGAT 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        884 CAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGAT 943

Qy        847 GCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAAC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 GCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAAC 1003

Qy        907 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 1063

Qy        967 CTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db       1064 CTTGCAGGACGAGATTTGATGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1123

Qy       1027 CTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1124 CTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG 1183

Qy       1087 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1184 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1243

Qy       1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1244 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1303

Qy       1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1304 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1363

Qy       1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1364 CC-GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1422

Qy       1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1482

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1542

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1602

Qy       1507 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1603 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1662

Qy       1567 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGA 1626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1663 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGA 1722

Qy       1627 ACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAA 1686
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       1723 ACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTATTGCAACTTTTCTTTGTCAA 1782

Qy       1687 GAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCT 1746
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db       1783 GAAAAAATATCAACTACAAGTATTCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCT 1842

Qy       1747 CTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGA 1806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1843 CTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGA 1902

Qy       1807 GGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGAT 1866
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db       1903 GGGCTGGATATTGAAAATGTGCAACATGTTATCAGTTTTGATCTTCCTTCTACCATTGAT 1962

Qy       1867 GAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCC 1926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1963 GAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCC 2022

Qy       1927 TTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGAT 1986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2023 TTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGAT 2082

Qy       1987 GCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGC 2046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2083 GCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGC 2142

Qy       2047 TTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGC 2106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2143 TTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGC 2202

Qy       2107 ACTTTGAACACAGCTGGGTTTTCTTCTTCAC 2137
              |||||||||||||||||||||||||||||||
Db       2203 ACTTTGAACACAGCTGGGTTTTCTTCTTCAC 2233
```

```
RESULT 2
BG718234
LOCUS       BG718234                 682 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602696209F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4828298 5',
            mRNA sequence.
ACCESSION   BG718234
VERSION     BG718234.1  GI:13997421
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10746  row: a  column: 03
            High quality sequence stop: 682.
FEATURES             Location/Qualifiers
     source          1..682
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4828298"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
```

```
                  Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      205 a    109 c    204 g    164 t
ORIGIN

  Query Match            28.4%;  Score 616;  DB 12;  Length 682;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 126

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 186

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 246

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 306

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 366

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 426

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 486

Qy        427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 546

Qy        487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 606

Qy        547 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 666

Qy        607 GGCAGTGGAAGTGAAC 622
              ||||||||||||||||
Db        667 GGCAGTGGAAGTGAAC 682


RESULT 3
BI560849
LOCUS       BI560849                 693 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603254018F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296444 5',
            mRNA sequence.
ACCESSION   BI560849
VERSION     BI560849.1  GI:15448163
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11750  row: c  column: 05
            High quality sequence stop: 691.
FEATURES             Location/Qualifiers
     source          1..693
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5296444"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      207 a    111 c    207 g    168 t
ORIGIN

  Query Match            28.4%;  Score 616;  DB 13;  Length 693;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 137

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 197

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 257

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 317

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 377

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 437

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 497

Qy        427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 557

Qy        487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 617

Qy        547 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 677

Qy        607 GGCAGTGGAAGTGAAC 622
              ||||||||||||||||
```

```
Db        678 GGCAGTGGAAGTGAAC 693

RESULT 4
BG717110
LOCUS       BG717110                798 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602689191F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821503 5',
            mRNA sequence.
ACCESSION   BG717110
VERSION     BG717110.1  GI:13996297
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 798)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10728  row: e  column: 24
            High quality sequence stop: 770.
FEATURES             Location/Qualifiers
     source          1..798
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4821503"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      254 a    124 c    233 g    186 t      1 others
ORIGIN

  Query Match             28.2%;  Score 612;  DB 12;  Length 798;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  662;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         16 TGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAGGATAGG 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAGGATAGG 125

Qy         76 TATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGAT 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGAT 185

Qy        136 GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTT 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTT 245

Qy        196 GGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTT 255
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GAAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTT 305

Qy        256 GGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 365

Qy        316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGA 425

Qy        376 GGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATAC 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 GGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATAC 485

Qy        436 AGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCA 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 AGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCA 545

Qy        496 AATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 AATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGA 605

Qy        556 AGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGA 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 AGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGA 665

Qy        616 AGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAAT 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 AGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAAT 725

Qy        676 TCT 678
              |||
Db        726 TCT 728


RESULT 5
BQ429126
LOCUS       BQ429126                792 bp    mRNA    linear   EST 24-MAY-2002
DEFINITION  AGENCOURT_7901787 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6104663
            5', mRNA sequence.
ACCESSION   BQ429126
VERSION     BQ429126.1  GI:21168202
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2342  row: n  column: 24
            High quality sequence stop: 496.
FEATURES             Location/Qualifiers
     source          1..792
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6104663"
                     /clone_lib="NIH_MGC_82"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).   Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
```

```
                    Alto, CA)."
BASE COUNT      241 a    128 c    242 g    180 t      1 others
ORIGIN

  Query Match             24.6%;  Score 535;  DB 14;  Length 792;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-281;
  Matches  535;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 126

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 186

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 246

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 306

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 366

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 426

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 486

Qy        427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 546

Qy        487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCC 541
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCC 601


RESULT 6
BG719449
LOCUS       BG719449                 747 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602690172F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822528 5',
            mRNA sequence.
ACCESSION   BG719449
VERSION     BG719449.1  GI:13998636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 747)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10730  row: p  column: 17
            High quality sequence stop: 734.
FEATURES             Location/Qualifiers
     source          1..747
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4822528"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      220 a    121 c    226 g    180 t
ORIGIN

  Query Match             23.4%;  Score 509;  DB 12;  Length 747;
  Best Local Similarity   99.8%;  Pred. No. 4.1e-267;
  Matches  629;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 162

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 222

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 282

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 342

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 402

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 462

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAA-ATAATTCAGAAGCTTC 425
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        463 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAACATAATTCAGAAGCTTC 522

Qy        426 AGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCT 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 AGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCT 582

Qy        486 AGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTT 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 AGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTT 642

Qy        546 TGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 TGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAG 702

Qy        606 TGGCAGTGGAAGTGAACGAGGTGGTTACAA 635
              ||||||||||||||||||||||||||||||
Db        703 TGGCAGTGGAAGTGAACGAGGTGGTTACAA 732


RESULT 7
BG717812
LOCUS       BG717812                 657 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602693829F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4826190 5',
            mRNA sequence.
```

```
ACCESSION   BG717812
VERSION     BG717812.1  GI:13996999
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10740  row: i  column: 07
            High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1..657
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4826190"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      200 a    105 c    192 g    160 t
ORIGIN

  Query Match             22.1%;  Score 480;  DB 12;  Length 657;
  Best Local Similarity   99.8%;  Pred. No. 3.2e-251;
  Matches  600;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 116

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 176

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 236

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 296

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 356

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 416

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCA-GAAGCTTC 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        417 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCACGAAGCTTC 476

Qy        426 AGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCT 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 AGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCT 536

Qy        486 AGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTT 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 AGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTT 596

Qy        546 TGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 TGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAG 656

Qy        606 T 606
              |
Db        657 T 657


RESULT 8
BG717354
LOCUS       BG717354                 651 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602689688F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822149 5',
            mRNA sequence.
ACCESSION   BG717354
VERSION     BG717354.1  GI:13996541
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 651)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10729  row: p  column: 22
            High quality sequence stop: 651.
FEATURES             Location/Qualifiers
     source          1..651
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4822149"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      194 a    105 c    195 g    157 t
ORIGIN

  Query Match             21.6%;  Score 469;  DB 12;  Length 651;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-245;
  Matches  469;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 TGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAGGATAGG 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 TGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAGGATAGG 136
```

```
Qy         76 TATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGAT 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 TATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGAT 196

Qy        136 GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTT 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTT 256

Qy        196 GGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTT 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 GGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTT 316

Qy        256 GGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 GGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 376

Qy        316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGA 436

Qy        376 GGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATAC 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATAC 496

Qy        436 AGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTC 484
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 AGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTC 545
```

```
RESULT 9
BI464267
LOCUS       BI464267                 822 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603203379F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269434 5',
            mRNA sequence.
ACCESSION   BI464267
VERSION     BI464267.1  GI:15254923
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11679  row: m  column: 19
            High quality sequence stop: 714.
FEATURES             Location/Qualifiers
     source          1..822
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5269434"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      255 a    140 c    239 g    188 t
ORIGIN

  Query Match             20.3%;  Score 441;  DB 13;  Length 822;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-230;
  Matches  441;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 129

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 189

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 249

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 309

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 369

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 429

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 489

Qy        427 GGGCCATACAGAAGAGGTGGA 447
              |||||||||||||||||||||
Db        490 GGGCCATACAGAAGAGGTGGA 510
```

```
RESULT 10
BG718634
LOCUS       BG718634                 846 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602696739F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4828977 5',
            mRNA sequence.
ACCESSION   BG718634
VERSION     BG718634.1  GI:13997821
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10747  row: m  column: 10
            High quality sequence stop: 701.
FEATURES             Location/Qualifiers
     source          1..846
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4828977"
```

```
                 /clone_lib="NIH_MGC_97"
                 /lab_host="DH10B"
                 /note="Organ: testis; Vector: pBluescriptR (modified
                 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                 size-selected for  average insert size 2.2 kb and
                 normalized to ROT 5. This is a primary library enriched
                 for full-length clones and constructed using the
                 Cap-trapper method (Carninci, in preparation).  Library
                 constructed by M. Brownstein (NIMH/NHGRI, National
                 Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      251 a    153 c    248 g    194 t
ORIGIN

  Query Match            20.3%;  Score 440;  DB 12;  Length 846;
  Best Local Similarity  100.0%;  Pred. No. 3e-229;
  Matches  440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 137

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 197

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 257

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 317

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 377

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 437

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 497

Qy        427 GGGCCATACAGAAGAGGTGG 446
              ||||||||||||||||||||
Db        498 GGGCCATACAGAAGAGGTGG 517


RESULT 11
BI461393
LOCUS       BI461393                 861 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603206861F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272430 5',
            mRNA sequence.
ACCESSION   BI461393
VERSION     BI461393.1  GI:15252049
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 861)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11687  row: j  column: 15
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1..861
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5272430"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      265 a    147 c    247 g    201 t      1 others
ORIGIN

  Query Match            19.8%;  Score 430;  DB 13;  Length 861;
  Best Local Similarity  99.7%;  Pred. No. 9.3e-224;
  Matches  600;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 114

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 174

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 234

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 294

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GGTGGTTNTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 354

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 414

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 474

Qy        427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        475 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGA-TTGGTCTA 533

Qy        487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 593

Qy        547 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 653

Qy        607 GG 608
              ||
Db        654 GG 655
```

```
RESULT 12
BI462207
LOCUS       BI462207                 830 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603205520F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5271221 5',
            mRNA sequence.
ACCESSION   BI462207
VERSION     BI462207.1  GI:15252863
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 830)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11684  row: h  column: 06
            High quality sequence stop: 769.
FEATURES             Location/Qualifiers
     source          1..830
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5271221"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      255 a    137 c    244 g    193 t      1 others
ORIGIN

  Query Match             19.4%;  Score 422;  DB 13;  Length 830;
  Best Local Similarity   99.8%;  Pred. No. 2.3e-219;
  Matches  472;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 137

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 197

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 257

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 317

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GGTGGTTNTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 377

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 437

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 497

Qy        427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATT 479
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATT 550


RESULT 13
BG771822
LOCUS       BG771822                 699 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602720504F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837296 5',
            mRNA sequence.
ACCESSION   BG771822
VERSION     BG771822.1  GI:14082475
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10769  row: h  column: 01
            High quality sequence stop: 692.
FEATURES             Location/Qualifiers
     source          1..699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4837296"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      208 a    113 c    211 g    167 t
ORIGIN

  Query Match             19.0%;  Score 413;  DB 12;  Length 699;
  Best Local Similarity   99.8%;  Pred. No. 1.8e-214;
  Matches  533;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 143

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 203

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        204 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 263

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 323

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 GGTGG-TTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 382

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 442

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 502

Qy        427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 562

Qy        487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 616
```

RESULT 14

```
BI463502
LOCUS       BI463502                743 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603203596F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269511 5',
            mRNA sequence.
ACCESSION   BI463502
VERSION     BI463502.1  GI:15254145
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11679  row: p  column: 24
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1..743
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5269511"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT    227 a    116 c    222 g    177 t      1 others
ORIGIN
```

```
  Query Match             19.0%;  Score 412;  DB 13;  Length 743;
  Best Local Similarity   99.8%;  Pred. No. 6.6e-214;
  Matches  462;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         17 GGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAGGATAGGT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 GGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAGGATAGGT 157

Qy         77 ATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGATG 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 ATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGATG 217

Qy        137 ATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTG 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTG 277

Qy        197 GAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTG 256
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        278 GAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTNG 337

Qy        257 GAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATA 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 GAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATA 397

Qy        317 GCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAG 457

Qy        377 GGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACA 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 GGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACA 517

Qy        437 GAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATT 479
              |||||||||||||||||||||||||||||||||||||||||||
Db        518 GAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATT 560
```

RESULT 15

```
BI459115
LOCUS       BI459115                841 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603199485F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5265501 5',
            mRNA sequence.
ACCESSION   BI459115
VERSION     BI459115.1  GI:15249771
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 841)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11669  row: i  column: 22
            High quality sequence stop: 646.
FEATURES             Location/Qualifiers
     source          1..841
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5265501"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
```

```
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      235 a    175 c    244 g    187 t
ORIGIN

  Query Match             18.0%;  Score 392;  DB 13;  Length 841;
  Best Local Similarity   99.8%;  Pred. No. 6.5e-203;
  Matches  512;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy         16 TGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAGGATAGG 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 TGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAGGATAGG 155

Qy         76 TATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGAT 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 TATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGAT 215

Qy        136 GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTT 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTT 275

Qy        196 GGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTT 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTT 335

Qy        256 GGAGTTGGAA-AGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGA 314
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 GGAGTTGGAACAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGA 395

Qy        315 TAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAG 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 TAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAG 455

Qy        375 AGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATA 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 AGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATA 515

Qy        435 CAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 CAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCC 575

Qy        495 AAATAATGACTTAGACCCAGACGAATGTATGCA 527
              |||||||||||||||||||||||||||||||||
Db        576 AAATAATGACTTAGACCCAGACGAATGTATGCA 608

Search completed: June 10, 2003, 14:51:37
Job time : 2833 secs
```

OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 14:05:25 ; Search time 296 Seconds
(without alignments)
10269.578 Million cell updates/sec

Title: US-09-714-865-15
Perfect score: 2172
Sequence: 1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 870385 seqs, 699768693 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	340	15.7	420	10	US-09-833-381-1475	Sequence 1475, Ap
	2	255	11.7	591	9	US-09-764-891-1170	Sequence 1170, Ap
	3	220	10.1	2194	9	US-09-764-891-7481	Sequence 7481, Ap
	4	220	10.1	2195	9	US-09-764-891-7482	Sequence 7482, Ap
	5	220	10.1	2197	9	US-09-764-891-7483	Sequence 7483, Ap
	6	21	1.0	1842	10	US-09-815-242-6893	Sequence 6893, Ap
	7	21	1.0	1830121	9	US-10-329-960-1	Sequence 1, Appli
	8	20	0.9	244	10	US-09-878-574-10286	Sequence 10286, A
	9	20	0.9	266	10	US-09-878-574-12453	Sequence 12453, A
c	10	19	0.9	138	10	US-09-864-761-28835	Sequence 28835, A
	11	19	0.9	418	9	US-09-918-995-35717	Sequence 35717, A
c	12	19	0.9	592	10	US-09-864-761-12255	Sequence 12255, A
c	13	19	0.9	1472	9	US-09-925-299-60	Sequence 60, Appl
c	14	19	0.9	1472	10	US-09-925-299-60	Sequence 60, Appl
	15	19	0.9	1987	10	US-09-925-301-185	Sequence 185, App
	16	19	0.9	2451	12	US-10-044-090-480	Sequence 480, App
c	17	19	0.9	4250	9	US-09-957-187-84	Sequence 84, Appl
	18	19	0.9	4423	12	US-10-002-600-67	Sequence 67, Appl
	19	19	0.9	536165	9	US-09-939-964-1	Sequence 1, Appli
	20	18	0.8	297	10	US-09-923-876-3712	Sequence 3712, Ap
c	21	18	0.8	379	9	US-09-991-936-1246	Sequence 1246, Ap
c	22	18	0.8	381	9	US-09-991-936-725	Sequence 725, App
c	23	18	0.8	413	9	US-09-918-995-7261	Sequence 7261, Ap
	24	18	0.8	421	9	US-10-076-622-116	Sequence 116, App
	25	18	0.8	421	10	US-09-604-287A-116	Sequence 116, App
	26	18	0.8	421	10	US-09-339-338-116	Sequence 116, App
	27	18	0.8	421	12	US-10-007-805-116	Sequence 116, App
	28	18	0.8	423	10	US-09-864-761-11121	Sequence 11121, A
	29	18	0.8	455	9	US-09-918-995-11308	Sequence 11308, A
	30	18	0.8	478	9	US-09-918-995-20686	Sequence 20686, A
	31	18	0.8	511	10	US-09-783-590-183	Sequence 183, App
	32	18	0.8	532	10	US-09-755-325-3	Sequence 3, Appli
c	33	18	0.8	587	9	US-09-232-785-336	Sequence 336, App
c	34	18	0.8	598	10	US-09-833-381-99	Sequence 99, Appl
c	35	18	0.8	599	9	US-09-954-531-816	Sequence 816, App
	36	18	0.8	606	9	US-10-198-846-7904	Sequence 7904, Ap
c	37	18	0.8	1205	9	US-09-533-029-5	Sequence 5, Appli
	38	18	0.8	1536	10	US-09-755-325-5	Sequence 5, Appli
	39	18	0.8	1614	9	US-09-938-842A-2560	Sequence 2560, Ap
	40	18	0.8	1682	10	US-09-755-325-7	Sequence 7, Appli
	41	18	0.8	1914	9	US-10-051-902-19	Sequence 19, Appl
	42	18	0.8	1914	9	US-10-051-909-19	Sequence 19, Appl
	43	18	0.8	2000	9	US-09-938-842A-3617	Sequence 3617, Ap
c	44	18	0.8	2005	12	US-10-142-373-1	Sequence 1, Appli
c	45	18	0.8	2196	9	US-10-086-464-16	Sequence 16, Appl
c	46	18	0.8	2261	9	US-10-086-464-15	Sequence 15, Appl
c	47	18	0.8	2715	9	US-09-950-041-5	Sequence 5, Appli
	48	18	0.8	3103	8	US-08-825-486-3	Sequence 3, Appli
	49	18	0.8	3103	8	US-08-870-434-2	Sequence 2, Appli
	50	18	0.8	3103	9	US-09-560-150-2	Sequence 2, Appli
	51	18	0.8	3103	9	US-10-067-741-2	Sequence 2, Appli
	52	18	0.8	3103	10	US-09-372-044-3	Sequence 3, Appli
	53	18	0.8	3103	10	US-09-924-417-66	Sequence 66, Appl
	54	18	0.8	3111	9	US-10-181-603-3	Sequence 3, Appli
	55	18	0.8	3111	10	US-09-954-456-1877	Sequence 1877, Ap
c	56	18	0.8	9551	9	US-09-764-872-825	Sequence 825, App
	57	18	0.8	13467	9	US-09-764-868-1330	Sequence 1330, Ap
	58	18	0.8	13467	9	US-09-764-868-1498	Sequence 1498, Ap
	59	18	0.8	13467	9	US-09-764-891-8632	Sequence 8632, Ap
c	60	18	0.8	15558	9	US-09-764-872-823	Sequence 823, App
c	61	18	0.8	18385	9	US-10-074-095-1018	Sequence 1018, Ap
c	62	18	0.8	18385	10	US-09-764-860-1018	Sequence 1018, Ap
	63	18	0.8	20020	9	US-10-091-483-325	Sequence 325, App
c	64	18	0.8	20020	9	US-09-764-891-8223	Sequence 8223, Ap
	65	18	0.8	20020	10	US-09-764-846-325	Sequence 325, App
	66	17	0.8	24	9	US-09-940-185-902	Sequence 902, App
c	67	17	0.8	37	9	US-09-780-533A-4312	Sequence 4312, Ap
	68	17	0.8	54	9	US-09-747-377-189	Sequence 189, App
	69	17	0.8	54	9	US-10-105-613-189	Sequence 189, App
c	70	17	0.8	135	10	US-09-864-761-31834	Sequence 31834, A
	71	17	0.8	155	10	US-09-864-761-27937	Sequence 27937, A
c	72	17	0.8	168	10	US-09-864-761-19712	Sequence 19712, A
	73	17	0.8	178	10	US-09-878-574-9249	Sequence 9249, Ap
c	74	17	0.8	201	10	US-09-878-574-8481	Sequence 8481, Ap
	75	17	0.8	275	10	US-09-728-446-581	Sequence 581, App
c	76	17	0.8	278	10	US-09-294-093B-4334	Sequence 4334, Ap
	77	17	0.8	308	9	US-10-040-739-408	Sequence 408, App
c	78	17	0.8	318	9	US-09-738-626-1436	Sequence 1436, Ap
c	79	17	0.8	320	9	US-09-764-872-199	Sequence 199, App
	80	17	0.8	370	10	US-09-770-791-361	Sequence 361, App
	81	17	0.8	384	9	US-09-918-995-37707	Sequence 37707, A
	82	17	0.8	392	10	US-09-960-352-9239	Sequence 9239, Ap
c	83	17	0.8	395	9	US-10-066-543-1610	Sequence 1610, Ap
c	84	17	0.8	421	10	US-09-867-701-6299	Sequence 6299, Ap
c	85	17	0.8	422	10	US-09-864-761-2931	Sequence 2931, Ap
c	86	17	0.8	424	10	US-09-878-574-5113	Sequence 5113, Ap
c	87	17	0.8	434	10	US-09-960-352-8795	Sequence 8795, Ap
c	88	17	0.8	450	10	US-09-960-352-9252	Sequence 9252, Ap
c	89	17	0.8	464	10	US-09-867-701-10666	Sequence 10666, A
	90	17	0.8	471	10	US-09-864-761-2245	Sequence 2245, Ap
c	91	17	0.8	472	10	US-09-864-761-310	Sequence 310, App
c	92	17	0.8	472	10	US-09-864-761-14159	Sequence 14159, A

	93	17	0.8	476	10	US-09-864-761-11357	Sequence 11357, A
	94	17	0.8	479	9	US-09-918-995-34562	Sequence 34562, A
c	95	17	0.8	489	9	US-10-046-935-7	Sequence 7, Appli
c	96	17	0.8	489	9	US-09-878-178-7	Sequence 7, Appli
c	97	17	0.8	489	9	US-10-146-502-7	Sequence 7, Appli
	98	17	0.8	490	10	US-09-867-550-2007	Sequence 2007, Ap
	99	17	0.8	500	10	US-09-827-998-52	Sequence 52, Appl
c	100	17	0.8	502	10	US-09-783-590-238	Sequence 238, App
c	101	17	0.8	512	10	US-09-864-761-15312	Sequence 15312, A
c	102	17	0.8	518	9	US-10-066-543-1729	Sequence 1729, Ap
c	103	17	0.8	537	9	US-09-972-473-12	Sequence 12, Appl
	104	17	0.8	571	9	US-10-046-935-171	Sequence 171, App
	105	17	0.8	571	9	US-09-878-178-171	Sequence 171, App
	106	17	0.8	571	9	US-10-146-502-171	Sequence 171, App
	107	17	0.8	593	9	US-10-066-543-235	Sequence 235, App
c	108	17	0.8	624	9	US-09-976-736-6	Sequence 6, Appli
	109	17	0.8	697	9	US-09-774-639-46	Sequence 46, Appl
	110	17	0.8	697	9	US-09-969-730-32	Sequence 32, Appl
c	111	17	0.8	702	9	US-09-972-473-10	Sequence 10, Appl
c	112	17	0.8	720	9	US-09-764-891-6426	Sequence 6426, Ap
	113	17	0.8	744	10	US-09-815-242-9363	Sequence 9363, Ap
	114	17	0.8	744	10	US-09-815-242-9524	Sequence 9524, Ap
c	115	17	0.8	780	9	US-09-976-736-5	Sequence 5, Appli
c	116	17	0.8	789	9	US-09-972-473-22	Sequence 22, Appl
c	117	17	0.8	828	9	US-09-764-872-851	Sequence 851, App
c	118	17	0.8	828	9	US-09-764-872-852	Sequence 852, App
c	119	17	0.8	828	9	US-09-764-872-853	Sequence 853, App
c	120	17	0.8	828	9	US-09-764-872-854	Sequence 854, App
c	121	17	0.8	884	9	US-10-002-344A-87	Sequence 87, Appl
c	122	17	0.8	915	9	US-10-115-928-26	Sequence 26, Appl
c	123	17	0.8	931	9	US-09-938-842A-2975	Sequence 2975, Ap
c	124	17	0.8	948	9	US-09-968-436B-5	Sequence 5, Appli
	125	17	0.8	987	10	US-09-974-300-1852	Sequence 1852, Ap
	126	17	0.8	1182	9	US-09-738-626-773	Sequence 773, App
	127	17	0.8	1206	9	US-09-738-626-1966	Sequence 1966, Ap
	128	17	0.8	1239	9	US-09-938-842A-1874	Sequence 1874, Ap
	129	17	0.8	1401	9	US-09-935-371-32	Sequence 32, Appl
c	130	17	0.8	1446	10	US-09-884-260A-8	Sequence 8, Appli
	131	17	0.8	1485	9	US-09-935-371-33	Sequence 33, Appl
c	132	17	0.8	1506	10	US-09-796-766-9	Sequence 9, Appli
c	133	17	0.8	1517	10	US-09-796-766-7	Sequence 7, Appli
c	134	17	0.8	1521	10	US-09-801-368-89	Sequence 89, Appl
c	135	17	0.8	1542	10	US-09-887-576-230	Sequence 230, App
	136	17	0.8	1674	9	US-09-764-891-8862	Sequence 8862, Ap
	137	17	0.8	1674	9	US-09-764-891-8865	Sequence 8865, Ap
c	138	17	0.8	1715	10	US-09-884-260A-56	Sequence 56, Appl
	139	17	0.8	1869	9	US-09-935-371-40	Sequence 40, Appl
c	140	17	0.8	1906	9	US-10-218-547-15	Sequence 15, Appl
c	141	17	0.8	1906	9	US-10-272-411-10	Sequence 10, Appl
	142	17	0.8	1924	10	US-09-765-272-159	Sequence 159, App
	143	17	0.8	2000	9	US-09-938-842A-4042	Sequence 4042, Ap
c	144	17	0.8	2000	9	US-09-938-842A-4108	Sequence 4108, Ap
	145	17	0.8	2000	9	US-09-938-842A-4329	Sequence 4329, Ap
c	146	17	0.8	2000	9	US-09-938-842A-5152	Sequence 5152, Ap
c	147	17	0.8	2120	9	US-09-344-882-29	Sequence 29, Appl
	148	17	0.8	2219	9	US-09-961-721-1	Sequence 1, Appli
	149	17	0.8	2235	9	US-09-769-787-196	Sequence 196, App
c	150	17	0.8	2509	9	US-10-108-605-86	Sequence 86, Appl
	151	17	0.8	2548	9	US-09-935-371-41	Sequence 41, Appl
c	152	17	0.8	2693	9	US-09-850-948-1	Sequence 1, Appli
	153	17	0.8	3124	9	US-10-219-220-131	Sequence 131, App
c	154	17	0.8	3170	9	US-10-174-590-249	Sequence 249, App
c	155	17	0.8	3170	9	US-10-176-758-249	Sequence 249, App
c	156	17	0.8	3170	9	US-10-175-737-249	Sequence 249, App
c	157	17	0.8	3170	9	US-10-173-706-249	Sequence 249, App
c	158	17	0.8	3170	9	US-10-175-738-249	Sequence 249, App
c	159	17	0.8	3170	9	US-10-175-752-249	Sequence 249, App
c	160	17	0.8	3170	9	US-10-176-482-249	Sequence 249, App
c	161	17	0.8	3170	9	US-10-176-757-249	Sequence 249, App
c	162	17	0.8	3170	9	US-10-176-913-249	Sequence 249, App
c	163	17	0.8	3170	9	US-10-180-552-249	Sequence 249, App
c	164	17	0.8	3170	9	US-10-180-557-249	Sequence 249, App
c	165	17	0.8	3170	9	US-10-173-700-249	Sequence 249, App
c	166	17	0.8	3170	9	US-10-174-572-249	Sequence 249, App
c	167	17	0.8	3170	9	US-10-174-579-249	Sequence 249, App
c	168	17	0.8	3170	9	US-10-174-582-249	Sequence 249, App
c	169	17	0.8	3170	9	US-10-174-588-249	Sequence 249, App
c	170	17	0.8	3170	9	US-10-175-739-249	Sequence 249, App
c	171	17	0.8	3170	9	US-10-175-740-249	Sequence 249, App
c	172	17	0.8	3170	9	US-10-175-743-249	Sequence 249, App
c	173	17	0.8	3170	9	US-10-176-488-249	Sequence 249, App
c	174	17	0.8	3170	9	US-10-176-492-249	Sequence 249, App
c	175	17	0.8	3170	9	US-10-176-747-249	Sequence 249, App
c	176	17	0.8	3170	9	US-10-176-750-249	Sequence 249, App
c	177	17	0.8	3170	9	US-10-176-985-249	Sequence 249, App
c	178	17	0.8	3170	9	US-10-176-987-249	Sequence 249, App
c	179	17	0.8	3170	9	US-10-176-991-249	Sequence 249, App
c	180	17	0.8	3170	9	US-10-176-992-249	Sequence 249, App
c	181	17	0.8	3170	9	US-10-176-993-249	Sequence 249, App
c	182	17	0.8	3170	9	US-10-184-658-249	Sequence 249, App
c	183	17	0.8	3170	9	US-10-173-695-249	Sequence 249, App
c	184	17	0.8	3170	9	US-10-173-697-249	Sequence 249, App
c	185	17	0.8	3170	9	US-10-173-705-249	Sequence 249, App
c	186	17	0.8	3170	9	US-10-174-576-249	Sequence 249, App
c	187	17	0.8	3170	9	US-10-174-585-249	Sequence 249, App
c	188	17	0.8	3170	9	US-10-174-586-249	Sequence 249, App
c	189	17	0.8	3170	9	US-10-175-747-249	Sequence 249, App
c	190	17	0.8	3170	9	US-10-176-481-249	Sequence 249, App
c	191	17	0.8	3170	9	US-10-176-485-249	Sequence 249, App
c	192	17	0.8	3170	9	US-10-176-487-249	Sequence 249, App
c	193	17	0.8	3170	9	US-10-176-493-249	Sequence 249, App
c	194	17	0.8	3170	9	US-10-176-756-249	Sequence 249, App
c	195	17	0.8	3170	9	US-10-176-911-249	Sequence 249, App
c	196	17	0.8	3170	9	US-10-176-919-249	Sequence 249, App
c	197	17	0.8	3170	9	US-10-176-925-249	Sequence 249, App
c	198	17	0.8	3170	9	US-10-176-978-249	Sequence 249, App
c	199	17	0.8	3170	9	US-10-179-510-249	Sequence 249, App
c	200	17	0.8	3170	9	US-10-180-543-249	Sequence 249, App
c	201	17	0.8	3170	9	US-10-180-544-249	Sequence 249, App
c	202	17	0.8	3170	9	US-10-180-546-249	Sequence 249, App
c	203	17	0.8	3170	9	US-10-180-547-249	Sequence 249, App
c	204	17	0.8	3170	9	US-10-180-549-249	Sequence 249, App
c	205	17	0.8	3170	9	US-10-180-555-249	Sequence 249, App
c	206	17	0.8	3170	9	US-10-180-559-249	Sequence 249, App
c	207	17	0.8	3170	9	US-10-181-000-249	Sequence 249, App
c	208	17	0.8	3170	9	US-10-183-010-249	Sequence 249, App
c	209	17	0.8	3170	9	US-10-183-012-249	Sequence 249, App
c	210	17	0.8	3170	9	US-10-184-614-249	Sequence 249, App
c	211	17	0.8	3170	9	US-10-184-623-249	Sequence 249, App
c	212	17	0.8	3170	9	US-10-184-635-249	Sequence 249, App
c	213	17	0.8	3170	9	US-10-184-637-249	Sequence 249, App
c	214	17	0.8	3170	9	US-10-184-646-249	Sequence 249, App
c	215	17	0.8	3170	9	US-10-184-647-249	Sequence 249, App
c	216	17	0.8	3170	9	US-10-184-652-249	Sequence 249, App
c	217	17	0.8	3170	9	US-10-187-594-249	Sequence 249, App
c	218	17	0.8	3170	9	US-10-187-596-249	Sequence 249, App
c	219	17	0.8	3170	9	US-10-187-745-249	Sequence 249, App
c	220	17	0.8	3170	9	US-10-187-885-249	Sequence 249, App
c	221	17	0.8	3170	9	US-10-187-886-249	Sequence 249, App
c	222	17	0.8	3170	9	US-10-199-464-249	Sequence 249, App
c	223	17	0.8	3170	9	US-10-176-751-249	Sequence 249, App
c	224	17	0.8	3170	9	US-10-176-760-249	Sequence 249, App
c	225	17	0.8	3170	9	US-10-176-990-249	Sequence 249, App
c	226	17	0.8	3170	9	US-10-180-541-249	Sequence 249, App
c	227	17	0.8	3170	9	US-10-180-542-249	Sequence 249, App
c	228	17	0.8	3170	9	US-10-180-548-249	Sequence 249, App
c	229	17	0.8	3170	9	US-10-180-551-249	Sequence 249, App
c	230	17	0.8	3170	9	US-10-180-998-249	Sequence 249, App
c	231	17	0.8	3170	9	US-10-180-999-249	Sequence 249, App
c	232	17	0.8	3170	9	US-10-183-013-249	Sequence 249, App
c	233	17	0.8	3170	9	US-10-184-612-249	Sequence 249, App
c	234	17	0.8	3170	9	US-10-184-616-249	Sequence 249, App
c	235	17	0.8	3170	9	US-10-184-617-249	Sequence 249, App
c	236	17	0.8	3170	9	US-10-184-622-249	Sequence 249, App
c	237	17	0.8	3170	9	US-10-184-628-249	Sequence 249, App
c	238	17	0.8	3170	9	US-10-184-629-249	Sequence 249, App

c 239	17	0.8	3170	9	US-10-184-630-249	Sequence 249, App
c 240	17	0.8	3170	9	US-10-184-631-249	Sequence 249, App
c 241	17	0.8	3170	9	US-10-184-632-249	Sequence 249, App
c 242	17	0.8	3170	9	US-10-184-636-249	Sequence 249, App
c 243	17	0.8	3170	9	US-10-184-640-249	Sequence 249, App
c 244	17	0.8	3170	9	US-10-184-650-249	Sequence 249, App
c 245	17	0.8	3170	9	US-10-184-651-249	Sequence 249, App
c 246	17	0.8	3170	9	US-10-187-588-249	Sequence 249, App
c 247	17	0.8	3170	9	US-10-187-597-249	Sequence 249, App
c 248	17	0.8	3170	9	US-10-187-598-249	Sequence 249, App
c 249	17	0.8	3170	9	US-10-187-600-249	Sequence 249, App
c 250	17	0.8	3170	9	US-10-187-601-249	Sequence 249, App
c 251	17	0.8	3170	9	US-10-187-602-249	Sequence 249, App
c 252	17	0.8	3170	9	US-10-187-603-249	Sequence 249, App
c 253	17	0.8	3170	9	US-10-187-741-249	Sequence 249, App
c 254	17	0.8	3170	9	US-10-187-743-249	Sequence 249, App
c 255	17	0.8	3170	9	US-10-187-746-249	Sequence 249, App
c 256	17	0.8	3170	9	US-10-187-747-249	Sequence 249, App
c 257	17	0.8	3170	9	US-10-187-751-249	Sequence 249, App
c 258	17	0.8	3170	9	US-10-187-753-249	Sequence 249, App
c 259	17	0.8	3170	9	US-10-187-754-249	Sequence 249, App
c 260	17	0.8	3170	9	US-10-187-757-249	Sequence 249, App
c 261	17	0.8	3170	9	US-10-187-884-249	Sequence 249, App
c 262	17	0.8	3170	9	US-10-188-767-249	Sequence 249, App
c 263	17	0.8	3170	9	US-10-188-769-249	Sequence 249, App
c 264	17	0.8	3170	9	US-10-188-770-249	Sequence 249, App
c 265	17	0.8	3170	9	US-10-188-773-249	Sequence 249, App
c 266	17	0.8	3170	9	US-10-188-781-249	Sequence 249, App
c 267	17	0.8	3170	9	US-10-194-361-249	Sequence 249, App
c 268	17	0.8	3170	9	US-10-194-423-249	Sequence 249, App
c 269	17	0.8	3170	9	US-10-195-897-249	Sequence 249, App
c 270	17	0.8	3170	9	US-10-195-901-249	Sequence 249, App
c 271	17	0.8	3170	9	US-10-196-756-249	Sequence 249, App
c 272	17	0.8	3170	9	US-10-173-708-249	Sequence 249, App
c 273	17	0.8	3170	9	US-10-176-479-249	Sequence 249, App
c 274	17	0.8	3170	9	US-10-176-748-249	Sequence 249, App
c 275	17	0.8	3170	9	US-10-176-916-249	Sequence 249, App
c 276	17	0.8	3170	9	US-10-179-507-249	Sequence 249, App
c 277	17	0.8	3170	9	US-10-179-516-249	Sequence 249, App
c 278	17	0.8	3170	9	US-10-179-519-249	Sequence 249, App
c 279	17	0.8	3170	9	US-10-179-525-249	Sequence 249, App
c 280	17	0.8	3170	9	US-10-180-540-249	Sequence 249, App
c 281	17	0.8	3170	9	US-10-180-545-249	Sequence 249, App
c 282	17	0.8	3170	9	US-10-183-006-249	Sequence 249, App
c 283	17	0.8	3170	9	US-10-183-008-249	Sequence 249, App
c 284	17	0.8	3170	9	US-10-183-017-249	Sequence 249, App
c 285	17	0.8	3170	9	US-10-183-019-249	Sequence 249, App
c 286	17	0.8	3170	9	US-10-184-618-249	Sequence 249, App
c 287	17	0.8	3170	9	US-10-184-625-249	Sequence 249, App
c 288	17	0.8	3170	9	US-10-184-626-249	Sequence 249, App
c 289	17	0.8	3170	9	US-10-184-627-249	Sequence 249, App
c 290	17	0.8	3170	9	US-10-184-645-249	Sequence 249, App
c 291	17	0.8	3170	9	US-10-184-654-249	Sequence 249, App
c 292	17	0.8	3170	9	US-10-184-655-249	Sequence 249, App
c 293	17	0.8	3170	9	US-10-188-774-249	Sequence 249, App
c 294	17	0.8	3170	9	US-10-188-775-249	Sequence 249, App
c 295	17	0.8	3170	9	US-10-194-462-249	Sequence 249, App
c 296	17	0.8	3170	9	US-10-195-902-249	Sequence 249, App
c 297	17	0.8	3170	9	US-10-196-743-249	Sequence 249, App
c 298	17	0.8	3170	9	US-10-196-745-249	Sequence 249, App
c 299	17	0.8	3170	9	US-10-196-760-249	Sequence 249, App
c 300	17	0.8	3170	9	US-10-196-762-249	Sequence 249, App
c 301	17	0.8	3170	9	US-10-197-695-249	Sequence 249, App
c 302	17	0.8	3170	9	US-10-006-856A-69	Sequence 69, Appl
c 303	17	0.8	3170	9	US-10-176-484-249	Sequence 249, App
c 304	17	0.8	3170	9	US-10-176-753-249	Sequence 249, App
c 305	17	0.8	3170	9	US-10-176-917-249	Sequence 249, App
c 306	17	0.8	3170	9	US-10-176-982-249	Sequence 249, App
c 307	17	0.8	3170	9	US-10-179-506-249	Sequence 249, App
c 308	17	0.8	3170	9	US-10-179-513-249	Sequence 249, App
c 309	17	0.8	3170	9	US-10-179-514-249	Sequence 249, App
c 310	17	0.8	3170	9	US-10-179-522-249	Sequence 249, App
c 311	17	0.8	3170	9	US-10-180-556-249	Sequence 249, App
c 312	17	0.8	3170	9	US-10-180-560-249	Sequence 249, App
c 313	17	0.8	3170	9	US-10-183-015-249	Sequence 249, App
c 314	17	0.8	3170	9	US-10-184-615-249	Sequence 249, App
c 315	17	0.8	3170	9	US-10-184-620-249	Sequence 249, App
c 316	17	0.8	3170	9	US-10-184-643-249	Sequence 249, App
c 317	17	0.8	3170	9	US-10-184-656-249	Sequence 249, App
c 318	17	0.8	3170	9	US-10-192-010-249	Sequence 249, App
c 319	17	0.8	3170	9	US-10-195-894-249	Sequence 249, App
c 320	17	0.8	3170	9	US-10-205-908-249	Sequence 249, App
c 321	17	0.8	3170	9	US-10-186-855-249	Sequence 249, App
c 322	17	0.8	3170	9	US-10-187-599-249	Sequence 249, App
c 323	17	0.8	3170	9	US-10-187-750-249	Sequence 249, App
c 324	17	0.8	3170	9	US-10-188-780-249	Sequence 249, App
c 325	17	0.8	3170	9	US-10-192-015-249	Sequence 249, App
c 326	17	0.8	3170	9	US-10-194-394-249	Sequence 249, App
c 327	17	0.8	3170	9	US-10-194-425-249	Sequence 249, App
c 328	17	0.8	3170	9	US-10-194-485-249	Sequence 249, App
c 329	17	0.8	3170	9	US-10-195-885-249	Sequence 249, App
c 330	17	0.8	3170	9	US-10-195-890-249	Sequence 249, App
c 331	17	0.8	3170	9	US-10-195-899-249	Sequence 249, App
c 332	17	0.8	3170	9	US-10-196-748-249	Sequence 249, App
c 333	17	0.8	3170	9	US-10-196-750-249	Sequence 249, App
c 334	17	0.8	3170	9	US-10-197-699-249	Sequence 249, App
c 335	17	0.8	3170	9	US-10-197-700-249	Sequence 249, App
c 336	17	0.8	3170	9	US-10-197-705-249	Sequence 249, App
c 337	17	0.8	3170	9	US-10-197-708-249	Sequence 249, App
c 338	17	0.8	3170	9	US-10-198-764-249	Sequence 249, App
c 339	17	0.8	3170	9	US-10-198-765-249	Sequence 249, App
c 340	17	0.8	3170	9	US-10-198-768-249	Sequence 249, App
c 341	17	0.8	3170	9	US-10-198-769-249	Sequence 249, App
c 342	17	0.8	3170	9	US-10-199-305-249	Sequence 249, App
c 343	17	0.8	3170	9	US-10-199-306-249	Sequence 249, App
c 344	17	0.8	3170	9	US-10-199-310-249	Sequence 249, App
c 345	17	0.8	3170	9	US-10-199-311-249	Sequence 249, App
c 346	17	0.8	3170	9	US-10-199-314-249	Sequence 249, App
c 347	17	0.8	3170	9	US-10-199-317-249	Sequence 249, App
c 348	17	0.8	3170	9	US-10-199-665-249	Sequence 249, App
c 349	17	0.8	3170	9	US-10-199-666-249	Sequence 249, App
c 350	17	0.8	3170	9	US-10-199-669-249	Sequence 249, App
c 351	17	0.8	3170	9	US-10-201-534-249	Sequence 249, App
c 352	17	0.8	3170	9	US-10-201-770-249	Sequence 249, App
c 353	17	0.8	3170	9	US-10-201-855-249	Sequence 249, App
c 354	17	0.8	3170	9	US-10-201-856-249	Sequence 249, App
c 355	17	0.8	3170	9	US-10-202-469-249	Sequence 249, App
c 356	17	0.8	3170	9	US-10-202-470-249	Sequence 249, App
c 357	17	0.8	3170	9	US-10-202-476-249	Sequence 249, App
c 358	17	0.8	3170	9	US-10-202-934-249	Sequence 249, App
c 359	17	0.8	3170	9	US-10-202-935-249	Sequence 249, App
c 360	17	0.8	3170	9	US-10-202-936-249	Sequence 249, App
c 361	17	0.8	3170	9	US-10-202-939-249	Sequence 249, App
c 362	17	0.8	3170	9	US-10-205-504-249	Sequence 249, App
c 363	17	0.8	3170	9	US-10-205-509-249	Sequence 249, App
c 364	17	0.8	3170	9	US-10-205-895-249	Sequence 249, App
c 365	17	0.8	3170	9	US-10-205-899-249	Sequence 249, App
c 366	17	0.8	3170	9	US-10-205-900-249	Sequence 249, App
c 367	17	0.8	3170	9	US-10-205-909-249	Sequence 249, App
c 368	17	0.8	3170	9	US-10-006-818A-69	Sequence 69, Appl
c 369	17	0.8	3170	9	US-10-183-002-249	Sequence 249, App
c 370	17	0.8	3170	9	US-10-184-621-249	Sequence 249, App
c 371	17	0.8	3170	9	US-10-184-638-249	Sequence 249, App
c 372	17	0.8	3170	9	US-10-187-752-249	Sequence 249, App
c 373	17	0.8	3170	9	US-10-187-887-249	Sequence 249, App
c 374	17	0.8	3170	9	US-10-194-461-249	Sequence 249, App
c 375	17	0.8	3170	9	US-10-195-892-249	Sequence 249, App
c 376	17	0.8	3170	9	US-10-196-751-249	Sequence 249, App
c 377	17	0.8	3170	9	US-10-197-694-249	Sequence 249, App
c 378	17	0.8	3170	9	US-10-197-697-249	Sequence 249, App
c 379	17	0.8	3170	9	US-10-197-707-249	Sequence 249, App
c 380	17	0.8	3170	9	US-10-199-303-249	Sequence 249, App
c 381	17	0.8	3170	9	US-10-199-318-249	Sequence 249, App
c 382	17	0.8	3170	9	US-10-199-458-249	Sequence 249, App
c 383	17	0.8	3170	9	US-10-199-462-249	Sequence 249, App
c 384	17	0.8	3170	9	US-10-201-324-249	Sequence 249, App

c	385	17	0.8	3170	9 US-10-201-328-249	Sequence 249, App
c	386	17	0.8	3170	9 US-10-201-527-249	Sequence 249, App
c	387	17	0.8	3170	9 US-10-201-528-249	Sequence 249, App
c	388	17	0.8	3170	9 US-10-201-529-249	Sequence 249, App
c	389	17	0.8	3170	9 US-10-201-530-249	Sequence 249, App
c	390	17	0.8	3170	9 US-10-202-408-249	Sequence 249, App
c	391	17	0.8	3170	9 US-10-202-409-249	Sequence 249, App
c	392	17	0.8	3170	9 US-10-202-411-249	Sequence 249, App
c	393	17	0.8	3170	9 US-10-202-472-249	Sequence 249, App
c	394	17	0.8	3170	9 US-10-205-502-249	Sequence 249, App
c	395	17	0.8	3170	9 US-10-205-507-249	Sequence 249, App
c	396	17	0.8	3170	9 US-10-205-511-249	Sequence 249, App
c	397	17	0.8	3170	9 US-10-205-902-249	Sequence 249, App
c	398	17	0.8	3170	9 US-10-205-907-249	Sequence 249, App
c	399	17	0.8	3170	9 US-10-194-456-249	Sequence 249, App
c	400	17	0.8	3170	9 US-10-196-758-249	Sequence 249, App
c	401	17	0.8	3170	9 US-10-198-770-249	Sequence 249, App
c	402	17	0.8	3170	9 US-10-199-308-249	Sequence 249, App
c	403	17	0.8	3170	9 US-10-200-617-249	Sequence 249, App
c	404	17	0.8	3170	9 US-10-205-893-249	Sequence 249, App
c	405	17	0.8	3170	9 US-10-205-897-249	Sequence 249, App
c	406	17	0.8	3170	9 US-10-015-393A-69	Sequence 69, Appl
c	407	17	0.8	3170	9 US-10-174-571-249	Sequence 249, App
c	408	17	0.8	3170	9 US-10-176-746-249	Sequence 249, App
c	409	17	0.8	3170	9 US-10-176-923-249	Sequence 249, App
c	410	17	0.8	3170	9 US-10-183-011-249	Sequence 249, App
c	411	17	0.8	3170	9 US-10-184-633-249	Sequence 249, App
c	412	17	0.8	3170	9 US-10-184-639-249	Sequence 249, App
c	413	17	0.8	3170	9 US-10-187-742-249	Sequence 249, App
c	414	17	0.8	3170	9 US-10-187-748-249	Sequence 249, App
c	415	17	0.8	3170	9 US-10-188-766-249	Sequence 249, App
c	416	17	0.8	3170	9 US-10-188-771-249	Sequence 249, App
c	417	17	0.8	3170	9 US-10-192-006-249	Sequence 249, App
c	418	17	0.8	3170	9 US-10-192-008-249	Sequence 249, App
c	419	17	0.8	3170	9 US-10-192-009-249	Sequence 249, App
c	420	17	0.8	3170	9 US-10-192-012-249	Sequence 249, App
c	421	17	0.8	3170	9 US-10-192-014-249	Sequence 249, App
c	422	17	0.8	3170	9 US-10-192-016-249	Sequence 249, App
c	423	17	0.8	3170	9 US-10-194-362-249	Sequence 249, App
c	424	17	0.8	3170	9 US-10-194-364-249	Sequence 249, App
c	425	17	0.8	3170	9 US-10-194-395-249	Sequence 249, App
c	426	17	0.8	3170	9 US-10-194-424-249	Sequence 249, App
c	427	17	0.8	3170	9 US-10-194-458-249	Sequence 249, App
c	428	17	0.8	3170	9 US-10-194-459-249	Sequence 249, App
c	429	17	0.8	3170	9 US-10-194-488-249	Sequence 249, App
c	430	17	0.8	3170	9 US-10-195-886-249	Sequence 249, App
c	431	17	0.8	3170	9 US-10-195-891-249	Sequence 249, App
c	432	17	0.8	3170	9 US-10-196-746-249	Sequence 249, App
c	433	17	0.8	3170	9 US-10-196-752-249	Sequence 249, App
c	434	17	0.8	3170	9 US-10-196-753-249	Sequence 249, App
c	435	17	0.8	3170	9 US-10-196-754-249	Sequence 249, App
c	436	17	0.8	3170	9 US-10-196-761-249	Sequence 249, App
c	437	17	0.8	3170	9 US-10-197-692-249	Sequence 249, App
c	438	17	0.8	3170	9 US-10-197-693-249	Sequence 249, App
c	439	17	0.8	3170	9 US-10-197-696-249	Sequence 249, App
c	440	17	0.8	3170	9 US-10-197-698-249	Sequence 249, App
c	441	17	0.8	3170	9 US-10-197-703-249	Sequence 249, App
c	442	17	0.8	3170	9 US-10-197-711-249	Sequence 249, App
c	443	17	0.8	3170	9 US-10-198-757-249	Sequence 249, App
c	444	17	0.8	3170	9 US-10-198-761-249	Sequence 249, App
c	445	17	0.8	3170	9 US-10-198-762-249	Sequence 249, App
c	446	17	0.8	3170	9 US-10-198-763-249	Sequence 249, App
c	447	17	0.8	3170	9 US-10-198-767-249	Sequence 249, App
c	448	17	0.8	3170	9 US-10-199-301-249	Sequence 249, App
c	449	17	0.8	3170	9 US-10-199-307-249	Sequence 249, App
c	450	17	0.8	3170	9 US-10-199-312-249	Sequence 249, App
c	451	17	0.8	3170	9 US-10-199-315-249	Sequence 249, App
c	452	17	0.8	3170	9 US-10-199-316-249	Sequence 249, App
c	453	17	0.8	3170	9 US-10-199-457-249	Sequence 249, App
c	454	17	0.8	3170	9 US-10-199-459-249	Sequence 249, App
c	455	17	0.8	3170	9 US-10-199-460-249	Sequence 249, App
c	456	17	0.8	3170	9 US-10-199-461-249	Sequence 249, App
c	457	17	0.8	3170	9 US-10-199-667-249	Sequence 249, App
c	458	17	0.8	3170	9 US-10-199-673-249	Sequence 249, App
c	459	17	0.8	3170	9 US-10-201-321-249	Sequence 249, App
c	460	17	0.8	3170	9 US-10-201-322-249	Sequence 249, App
c	461	17	0.8	3170	9 US-10-201-326-249	Sequence 249, App
c	462	17	0.8	3170	9 US-10-201-532-249	Sequence 249, App
c	463	17	0.8	3170	9 US-10-201-533-249	Sequence 249, App
c	464	17	0.8	3170	9 US-10-201-535-249	Sequence 249, App
c	465	17	0.8	3170	9 US-10-201-769-249	Sequence 249, App
c	466	17	0.8	3170	9 US-10-201-771-249	Sequence 249, App
c	467	17	0.8	3170	9 US-10-201-854-249	Sequence 249, App
c	468	17	0.8	3170	9 US-10-202-410-249	Sequence 249, App
c	469	17	0.8	3170	9 US-10-202-473-249	Sequence 249, App
c	470	17	0.8	3170	9 US-10-202-474-249	Sequence 249, App
c	471	17	0.8	3170	9 US-10-205-503-249	Sequence 249, App
c	472	17	0.8	3170	9 US-10-205-512-249	Sequence 249, App
c	473	17	0.8	3170	9 US-10-205-892-249	Sequence 249, App
c	474	17	0.8	3170	9 US-10-205-894-249	Sequence 249, App
c	475	17	0.8	3170	9 US-10-205-896-249	Sequence 249, App
c	476	17	0.8	3170	9 US-10-205-898-249	Sequence 249, App
c	477	17	0.8	3170	9 US-10-205-901-249	Sequence 249, App
c	478	17	0.8	3170	9 US-10-205-903-249	Sequence 249, App
c	479	17	0.8	3170	9 US-10-206-909-249	Sequence 249, App
c	480	17	0.8	3170	9 US-10-206-910-249	Sequence 249, App
c	481	17	0.8	3170	9 US-10-206-911-249	Sequence 249, App
c	482	17	0.8	3170	9 US-10-206-912-249	Sequence 249, App
c	483	17	0.8	3170	9 US-10-206-913-249	Sequence 249, App
c	484	17	0.8	3170	9 US-10-206-914-249	Sequence 249, App
c	485	17	0.8	3170	9 US-10-206-920-249	Sequence 249, App
c	486	17	0.8	3170	9 US-10-206-921-249	Sequence 249, App
c	487	17	0.8	3170	9 US-10-206-923-249	Sequence 249, App
c	488	17	0.8	3170	9 US-10-206-925-249	Sequence 249, App
c	489	17	0.8	3170	9 US-10-206-926-249	Sequence 249, App
c	490	17	0.8	3170	9 US-10-206-927-249	Sequence 249, App
c	491	17	0.8	3170	9 US-10-207-916-249	Sequence 249, App
c	492	17	0.8	3170	9 US-10-207-917-249	Sequence 249, App
c	493	17	0.8	3170	9 US-10-207-918-249	Sequence 249, App
c	494	17	0.8	3170	9 US-10-207-919-249	Sequence 249, App
c	495	17	0.8	3170	9 US-10-207-920-249	Sequence 249, App
c	496	17	0.8	3170	9 US-10-207-925-249	Sequence 249, App
c	497	17	0.8	3170	9 US-10-208-021-249	Sequence 249, App
c	498	17	0.8	3170	9 US-10-208-022-249	Sequence 249, App
c	499	17	0.8	3170	9 US-10-208-023-249	Sequence 249, App
c	500	17	0.8	3170	9 US-10-208-026-249	Sequence 249, App
c	501	17	0.8	3170	9 US-10-208-029-249	Sequence 249, App
c	502	17	0.8	3170	9 US-10-208-030-249	Sequence 249, App
c	503	17	0.8	3170	9 US-10-232-232-249	Sequence 249, App
c	504	17	0.8	3170	9 US-09-946-374-69	Sequence 69, Appl
c	505	17	0.8	3170	9 US-10-012-121A-69	Sequence 69, Appl
c	506	17	0.8	3170	9 US-10-015-869A-69	Sequence 69, Appl
c	507	17	0.8	3170	9 US-10-173-693-249	Sequence 249, App
c	508	17	0.8	3170	9 US-10-174-578-249	Sequence 249, App
c	509	17	0.8	3170	9 US-10-175-741-249	Sequence 249, App
c	510	17	0.8	3170	9 US-10-175-750-249	Sequence 249, App
c	511	17	0.8	3170	9 US-10-176-986-249	Sequence 249, App
c	512	17	0.8	3170	9 US-10-184-641-249	Sequence 249, App
c	513	17	0.8	3170	9 US-10-187-888-249	Sequence 249, App
c	514	17	0.8	3170	9 US-10-194-360-249	Sequence 249, App
c	515	17	0.8	3170	9 US-10-194-365-249	Sequence 249, App
c	516	17	0.8	3170	9 US-10-195-895-249	Sequence 249, App
c	517	17	0.8	3170	9 US-10-195-898-249	Sequence 249, App
c	518	17	0.8	3170	9 US-10-196-759-249	Sequence 249, App
c	519	17	0.8	3170	9 US-10-199-302-249	Sequence 249, App
c	520	17	0.8	3170	9 US-10-201-323-249	Sequence 249, App
c	521	17	0.8	3170	9 US-10-205-510-249	Sequence 249, App
c	522	17	0.8	3170	9 US-10-205-891-249	Sequence 249, App
c	523	17	0.8	3170	9 US-10-205-904-249	Sequence 249, App
c	524	17	0.8	3170	9 US-10-206-917-249	Sequence 249, App
c	525	17	0.8	3170	9 US-10-207-923-249	Sequence 249, App
c	526	17	0.8	3170	9 US-10-207-924-249	Sequence 249, App
c	527	17	0.8	3170	9 US-10-208-028-249	Sequence 249, App
c	528	17	0.8	3170	9 US-10-121-062-249	Sequence 249, App
c	529	17	0.8	3170	9 US-10-175-753-249	Sequence 249, App
c	530	17	0.8	3170	9 US-10-180-553-249	Sequence 249, App

c 531	17	0.8	3170	9	US-10-201-327-249	Sequence 249, App
c 532	17	0.8	3170	9	US-10-006-116A-69	Sequence 69, Appl
c 533	17	0.8	3170	9	US-10-006-117A-69	Sequence 69, Appl
c 534	17	0.8	3170	9	US-10-013-913A-69	Sequence 69, Appl
c 535	17	0.8	3170	9	US-10-017-527A-69	Sequence 69, Appl
c 536	17	0.8	3170	9	US-10-173-696-249	Sequence 249, App
c 537	17	0.8	3170	9	US-10-183-003-249	Sequence 249, App
c 538	17	0.8	3170	9	US-10-183-016-249	Sequence 249, App
c 539	17	0.8	3170	9	US-10-125-923A-249	Sequence 249, App
c 540	17	0.8	3170	9	US-10-176-491-249	Sequence 249, App
c 541	17	0.8	3170	9	US-10-176-979-249	Sequence 249, App
c 542	17	0.8	3170	9	US-10-187-592-249	Sequence 249, App
c 543	17	0.8	3170	9	US-10-007-194A-69	Sequence 69, Appl
c 544	17	0.8	3170	9	US-10-013-430A-69	Sequence 69, Appl
c 545	17	0.8	3170	9	US-10-197-691-249	Sequence 249, App
c 546	17	0.8	3170	9	US-10-198-771-249	Sequence 249, App
c 547	17	0.8	3170	9	US-10-011-671A-69	Sequence 69, Appl
c 548	17	0.8	3170	9	US-10-012-755A-69	Sequence 69, Appl
c 549	17	0.8	3170	9	US-10-174-575A-249	Sequence 249, App
c 550	17	0.8	3170	9	US-10-179-520-249	Sequence 249, App
c 551	17	0.8	3170	9	US-10-201-325-249	Sequence 249, App
c 552	17	0.8	3170	9	US-10-202-941-249	Sequence 249, App
c 553	17	0.8	3170	9	US-10-205-910-249	Sequence 249, App
c 554	17	0.8	3170	9	US-10-015-386A-69	Sequence 69, Appl
c 555	17	0.8	3170	9	US-10-179-526-249	Sequence 249, App
c 556	17	0.8	3170	12	US-10-052-586-249	Sequence 249, App
c 557	17	0.8	3422	9	US-09-764-891-6046	Sequence 6046, Ap
c 558	17	0.8	3422	9	US-09-764-891-6048	Sequence 6048, Ap
c 559	17	0.8	3422	9	US-10-091-438-271	Sequence 271, App
560	17	0.8	3422	9	US-10-091-438-273	Sequence 273, App
c 561	17	0.8	3696	9	US-09-972-473-20	Sequence 20, Appl
c 562	17	0.8	3718	10	US-09-917-800A-1471	Sequence 1471, Ap
563	17	0.8	4086	9	US-10-161-803-18	Sequence 18, Appl
564	17	0.8	4189	9	US-10-161-803-19	Sequence 19, Appl
c 565	17	0.8	4315	12	US-10-044-090-526	Sequence 526, App
c 566	17	0.8	4359	10	US-09-764-877-2346	Sequence 2346, Ap
c 567	17	0.8	4360	10	US-09-764-877-2347	Sequence 2347, Ap
568	17	0.8	4465	9	US-09-953-067A-2	Sequence 2, Appli
569	17	0.8	4496	9	US-10-071-766-46	Sequence 46, Appl
570	17	0.8	4496	12	US-10-044-090-478	Sequence 478, App
c 571	17	0.8	4998	10	US-09-738-968-42	Sequence 42, Appl
572	17	0.8	5881	9	US-09-764-891-9918	Sequence 9918, Ap
c 573	17	0.8	6226	10	US-09-827-949-1	Sequence 1, Appli
574	17	0.8	7214	9	US-10-205-942-1	Sequence 1, Appli
c 575	17	0.8	8151	9	US-10-205-942-2	Sequence 2, Appli
576	17	0.8	9453	10	US-09-954-456-324	Sequence 324, App
c 577	17	0.8	10198	9	US-09-764-891-8743	Sequence 8743, Ap
c 578	17	0.8	10198	9	US-09-764-891-8744	Sequence 8744, Ap
579	17	0.8	11064	9	US-09-949-293-25	Sequence 25, Appl
580	17	0.8	11735	9	US-10-239-676-34	Sequence 34, Appl
581	17	0.8	14800	10	US-09-954-456-1601	Sequence 1601, Ap
582	17	0.8	18404	9	US-10-142-515-10	Sequence 10, Appl
583	17	0.8	31129	9	US-10-298-192-3	Sequence 3, Appli
c 584	17	0.8	32154	10	US-09-764-877-3433	Sequence 3433, Ap
c 585	17	0.8	32768	10	US-09-070-927A-17	Sequence 17, Appl
586	17	0.8	42571	9	US-10-224-413-3	Sequence 3, Appli
c 587	17	0.8	53226	10	US-09-818-264-3	Sequence 3, Appli
588	17	0.8	82938	9	US-09-818-657-3	Sequence 3, Appli
589	17	0.8	180216	10	US-09-835-232-6	Sequence 6, Appli
c 590	17	0.8	326014	10	US-09-731-231A-3	Sequence 3, Appli
c 591	17	0.8	397658	10	US-09-813-320-3	Sequence 3, Appli
592	17	0.8	513509	9	US-09-754-853A-4	Sequence 4, Appli
c 593	17	0.8	513509	9	US-09-754-853A-4	Sequence 4, Appli
c 594	17	0.8	1503841	9	US-09-946-807-1	Sequence 1, Appli
c 595	17	0.8	1503841	10	US-09-795-668-1	Sequence 1, Appli
c 596	17	0.8	1503841	10	US-09-795-686-1	Sequence 1, Appli
c 597	17	0.8	1830121	9	US-10-329-960-1	Sequence 1, Appli
598	17	0.8	3309400	9	US-09-738-626-1	Sequence 1, Appli
c 599	17	0.8	3309400	9	US-09-738-626-1	Sequence 1, Appli
600	16	0.7	112	10	US-09-864-761-28203	Sequence 28203, A
c 601	16	0.7	118	9	US-10-005-338B-117	Sequence 117, App
602	16	0.7	141	10	US-09-864-761-25877	Sequence 25877, A
c 603	16	0.7	155	10	US-09-864-761-22570	Sequence 22570, A
604	16	0.7	175	10	US-09-878-574-12932	Sequence 12932, A
605	16	0.7	177	9	US-10-083-357-267	Sequence 267, App
606	16	0.7	180	10	US-09-867-701-6800	Sequence 6800, Ap
607	16	0.7	192	10	US-09-783-590-162	Sequence 162, App
c 608	16	0.7	207	10	US-09-864-761-29411	Sequence 29411, A
c 609	16	0.7	224	10	US-09-876-889-105	Sequence 105, App
610	16	0.7	229	10	US-09-878-574-5644	Sequence 5644, Ap
c 611	16	0.7	234	9	US-10-025-380-209	Sequence 209, App
c 612	16	0.7	234	10	US-09-922-217-209	Sequence 209, App
c 613	16	0.7	234	10	US-09-833-263-209	Sequence 209, App
614	16	0.7	235	10	US-09-923-876-1435	Sequence 1435, Ap
615	16	0.7	236	10	US-09-924-035A-153	Sequence 153, App
c 616	16	0.7	237	9	US-09-938-842A-1450	Sequence 1450, Ap
c 617	16	0.7	240	9	US-10-083-357-265	Sequence 265, App
c 618	16	0.7	247	10	US-09-864-761-30341	Sequence 30341, A
c 619	16	0.7	250	9	US-09-736-457-27	Sequence 27, Appl
c 620	16	0.7	250	9	US-09-902-941-27	Sequence 27, Appl
c 621	16	0.7	250	9	US-09-849-626-27	Sequence 27, Appl
c 622	16	0.7	250	9	US-10-017-754-27	Sequence 27, Appl
623	16	0.7	254	10	US-09-880-107-1346	Sequence 1346, Ap
624	16	0.7	260	10	US-09-923-876-6003	Sequence 6003, Ap
625	16	0.7	265	10	US-09-923-876-738	Sequence 738, App
626	16	0.7	269	9	US-09-809-391-142	Sequence 142, App
c 627	16	0.7	270	10	US-09-833-790-83	Sequence 83, Appl
c 628	16	0.7	270	10	US-09-833-790-377	Sequence 377, App
c 629	16	0.7	273	10	US-09-878-574-12212	Sequence 12212, A
630	16	0.7	274	10	US-09-563-817-423	Sequence 423, App
631	16	0.7	277	10	US-09-878-574-15242	Sequence 15242, A
c 632	16	0.7	283	10	US-09-783-590-5853	Sequence 5853, Ap
633	16	0.7	285	10	US-09-294-093B-831	Sequence 831, App
634	16	0.7	286	10	US-09-878-574-14644	Sequence 14644, A
c 635	16	0.7	304	10	US-09-864-761-21581	Sequence 21581, A
c 636	16	0.7	327	10	US-09-815-242-2750	Sequence 2750, Ap
c 637	16	0.7	338	10	US-09-867-701-7627	Sequence 7627, Ap
638	16	0.7	343	10	US-09-983-965-2317	Sequence 2317, Ap
639	16	0.7	345	10	US-09-878-574-2619	Sequence 2619, Ap
640	16	0.7	350	10	US-09-745-288-19	Sequence 19, Appl
c 641	16	0.7	351	7	US-08-781-986A-2969	Sequence 2969, Ap
642	16	0.7	362	10	US-09-878-574-1744	Sequence 1744, Ap
c 643	16	0.7	364	9	US-10-015-219-665	Sequence 665, App
c 644	16	0.7	364	10	US-09-777-564-665	Sequence 665, App
c 645	16	0.7	366	9	US-10-015-219-792	Sequence 792, App
c 646	16	0.7	366	10	US-09-777-564-792	Sequence 792, App
c 647	16	0.7	367	10	US-09-925-301-78	Sequence 78, Appl
c 648	16	0.7	371	10	US-09-864-761-11731	Sequence 11731, A
649	16	0.7	372	9	US-09-918-995-37446	Sequence 37446, A
650	16	0.7	373	10	US-09-764-877-315	Sequence 315, App
c 651	16	0.7	377	9	US-09-918-995-29711	Sequence 29711, A
652	16	0.7	378	10	US-09-878-574-3365	Sequence 3365, Ap
653	16	0.7	380	9	US-09-918-995-7788	Sequence 7788, Ap
654	16	0.7	381	9	US-09-796-692-8999	Sequence 8999, Ap
655	16	0.7	381	9	US-10-040-862-8999	Sequence 8999, Ap
c 656	16	0.7	383	9	US-09-803-719-1392	Sequence 1392, Ap
c 657	16	0.7	390	9	US-09-938-842A-2688	Sequence 2688, Ap
658	16	0.7	393	10	US-09-867-701-3238	Sequence 3238, Ap
c 659	16	0.7	393	10	US-09-974-300-7553	Sequence 7553, Ap
c 660	16	0.7	394	10	US-09-833-381-2021	Sequence 2021, Ap
c 661	16	0.7	395	9	US-10-060-036-300	Sequence 300, App
662	16	0.7	397	10	US-09-960-352-10508	Sequence 10508, A
c 663	16	0.7	401	9	US-09-946-807-624	Sequence 624, App
c 664	16	0.7	401	9	US-09-946-807-1125	Sequence 1125, Ap
c 665	16	0.7	401	10	US-09-795-668-624	Sequence 624, App
c 666	16	0.7	401	10	US-09-795-668-1125	Sequence 1125, App
c 667	16	0.7	401	10	US-09-795-686-624	Sequence 624, App
c 668	16	0.7	401	10	US-09-795-686-1125	Sequence 1125, Ap
c 669	16	0.7	404	9	US-09-918-995-36717	Sequence 36717, A
670	16	0.7	405	10	US-09-960-352-6748	Sequence 6748, Ap
671	16	0.7	407	10	US-09-878-574-1001	Sequence 1001, Ap
c 672	16	0.7	407	10	US-09-878-574-5018	Sequence 5018, Ap
673	16	0.7	408	10	US-09-878-574-439	Sequence 439, App
674	16	0.7	409	10	US-09-880-107-971	Sequence 971, App
675	16	0.7	421	9	US-09-918-995-6945	Sequence 6945, Ap
676	16	0.7	421	9	US-09-918-995-36017	Sequence 36017, A

	677	16	0.7	424	10	US-09-560-863-249 Sequence 249, App
	678	16	0.7	429	10	US-09-998-598-279 Sequence 279, App
	679	16	0.7	438	10	US-09-960-352-8396 Sequence 8396, Ap
c	680	16	0.7	441	10	US-09-960-352-3512 Sequence 3512, Ap
	681	16	0.7	443	10	US-09-867-701-2506 Sequence 2506, Ap
	682	16	0.7	446	10	US-09-879-536-629 Sequence 629, App
	683	16	0.7	448	9	US-09-918-995-9655 Sequence 9655, Ap
	684	16	0.7	449	10	US-09-983-965-2802 Sequence 2802, Ap
	685	16	0.7	455	10	US-09-867-550-61 Sequence 61, Appl
	686	16	0.7	461	10	US-09-867-701-3626 Sequence 3626, Ap
	687	16	0.7	465	9	US-09-918-995-14180 Sequence 14180, A
	688	16	0.7	467	10	US-09-864-761-6179 Sequence 6179, Ap
	689	16	0.7	468	10	US-09-864-761-15621 Sequence 15621, A
c	690	16	0.7	472	10	US-09-864-761-14629 Sequence 14629, A
c	691	16	0.7	472	10	US-09-764-877-726 Sequence 726, App
	692	16	0.7	474	10	US-09-864-761-11618 Sequence 11618, A
c	693	16	0.7	476	7	US-08-781-986A-1469 Sequence 1469, Ap
c	694	16	0.7	477	9	US-09-918-995-9777 Sequence 9777, Ap
c	695	16	0.7	477	9	US-09-918-995-12297 Sequence 12297, A
	696	16	0.7	477	10	US-09-998-598-654 Sequence 654, App
	697	16	0.7	478	9	US-09-918-995-7269 Sequence 7269, Ap
c	698	16	0.7	479	9	US-09-918-995-35584 Sequence 35584, A
	699	16	0.7	484	9	US-09-796-692-7 Sequence 7, Appli
	700	16	0.7	484	9	US-09-796-692-3311 Sequence 3311, Ap
	701	16	0.7	484	9	US-09-918-995-19639 Sequence 19639, A
	702	16	0.7	484	9	US-10-040-862-7 Sequence 7, Appli
	703	16	0.7	484	9	US-10-040-862-3311 Sequence 3311, Ap
	704	16	0.7	487	9	US-10-015-219-1218 Sequence 1218, Ap
	705	16	0.7	487	10	US-09-777-564-1218 Sequence 1218, Ap
c	706	16	0.7	489	9	US-10-001-883-30 Sequence 30, Appl
	707	16	0.7	491	9	US-09-918-995-20779 Sequence 20779, A
	708	16	0.7	495	9	US-09-918-995-20203 Sequence 20203, A
	709	16	0.7	495	9	US-09-918-995-21514 Sequence 21514, A
	710	16	0.7	495	10	US-09-880-107-1140 Sequence 1140, Ap
	711	16	0.7	496	9	US-10-046-935-407 Sequence 407, App
	712	16	0.7	496	9	US-09-878-178-407 Sequence 407, App
	713	16	0.7	496	9	US-10-146-502-407 Sequence 407, App
c	714	16	0.7	497	10	US-09-864-761-10862 Sequence 10862, A
	715	16	0.7	501	9	US-09-938-842A-2038 Sequence 2038, Ap
	716	16	0.7	501	9	US-09-954-692-97 Sequence 97, Appl
	717	16	0.7	501	10	US-09-559-671A-97 Sequence 97, Appl
c	718	16	0.7	502	10	US-09-864-761-13439 Sequence 13439, A
	719	16	0.7	505	9	US-10-091-504-1792 Sequence 1792, Ap
	720	16	0.7	505	10	US-09-764-869-1792 Sequence 1792, Ap
c	721	16	0.7	508	9	US-10-060-036-1137 Sequence 1137, Ap
c	722	16	0.7	509	10	US-09-998-598-522 Sequence 522, App
	723	16	0.7	510	9	US-09-918-995-32343 Sequence 32343, A
c	724	16	0.7	516	9	US-10-015-219-944 Sequence 944, App
c	725	16	0.7	516	10	US-09-777-564-944 Sequence 944, App
	726	16	0.7	517	10	US-09-974-300-3195 Sequence 3195, Ap
c	727	16	0.7	518	10	US-09-864-761-7886 Sequence 7886, Ap
c	728	16	0.7	527	10	US-09-833-381-1083 Sequence 1083, Ap
c	729	16	0.7	530	10	US-09-867-701-5223 Sequence 5223, Ap
	730	16	0.7	533	10	US-09-867-701-5295 Sequence 5295, Ap
	731	16	0.7	541	9	US-10-091-504-1793 Sequence 1793, Ap
	732	16	0.7	541	10	US-09-764-869-1793 Sequence 1793, Ap
c	733	16	0.7	547	9	US-09-991-936-301 Sequence 301, App
c	734	16	0.7	549	10	US-09-864-761-12848 Sequence 12848, A
	735	16	0.7	558	10	US-09-998-598-586 Sequence 586, App
	736	16	0.7	565	10	US-09-827-669-56 Sequence 56, Appl
c	737	16	0.7	566	10	US-09-864-761-9834 Sequence 9834, Ap
c	738	16	0.7	573	10	US-09-998-598-1184 Sequence 1184, Ap
	739	16	0.7	587	10	US-09-864-761-9344 Sequence 9344, Ap
c	740	16	0.7	589	9	US-10-015-219-101 Sequence 101, App
c	741	16	0.7	589	10	US-09-777-564-101 Sequence 101, App
	742	16	0.7	602	10	US-09-998-598-1275 Sequence 1275, Ap
c	743	16	0.7	621	9	US-09-796-692-3953 Sequence 3953, Ap
c	744	16	0.7	621	9	US-10-040-862-3953 Sequence 3953, Ap
	745	16	0.7	625	10	US-09-770-149-748 Sequence 748, App
	746	16	0.7	659	10	US-09-881-752A-277 Sequence 277, App
c	747	16	0.7	661	10	US-09-822-849A-584 Sequence 584, App
	748	16	0.7	666	9	US-10-198-846-8598 Sequence 8598, Ap
	749	16	0.7	694	10	US-09-815-242-4776 Sequence 4776, Ap
c	750	16	0.7	702	9	US-10-091-504-2231 Sequence 2231, Ap
c	751	16	0.7	702	9	US-10-091-504-2232 Sequence 2232, Ap
c	752	16	0.7	702	10	US-09-764-869-2231 Sequence 2231, Ap
c	753	16	0.7	702	10	US-09-764-869-2232 Sequence 2232, Ap
c	754	16	0.7	703	10	US-09-910-943-302 Sequence 302, App
c	755	16	0.7	704	9	US-10-091-504-2233 Sequence 2233, Ap
c	756	16	0.7	704	10	US-09-764-869-2233 Sequence 2233, Ap
c	757	16	0.7	714	9	US-10-144-929-72 Sequence 72, Appl
	758	16	0.7	750	9	US-10-101-464A-242 Sequence 242, App
c	759	16	0.7	755	9	US-09-992-598-15 Sequence 15, Appl
c	760	16	0.7	755	9	US-09-989-293A-15 Sequence 15, Appl
c	761	16	0.7	755	9	US-09-989-735-15 Sequence 15, Appl
c	762	16	0.7	755	9	US-09-990-444-15 Sequence 15, Appl
c	763	16	0.7	755	9	US-09-989-730-15 Sequence 15, Appl
c	764	16	0.7	755	9	US-09-990-436-15 Sequence 15, Appl
c	765	16	0.7	755	9	US-09-991-181-15 Sequence 15, Appl
c	766	16	0.7	755	9	US-09-993-687-15 Sequence 15, Appl
c	767	16	0.7	755	9	US-09-989-734-15 Sequence 15, Appl
c	768	16	0.7	755	9	US-09-997-653-15 Sequence 15, Appl
c	769	16	0.7	755	9	US-09-993-667-15 Sequence 15, Appl
c	770	16	0.7	755	9	US-09-990-438-15 Sequence 15, Appl
c	771	16	0.7	755	9	US-09-990-562-15 Sequence 15, Appl
c	772	16	0.7	755	9	US-09-997-428-15 Sequence 15, Appl
c	773	16	0.7	755	9	US-09-997-666-15 Sequence 15, Appl
c	774	16	0.7	755	9	US-09-990-711-15 Sequence 15, Appl
c	775	16	0.7	755	9	US-09-989-726-15 Sequence 15, Appl
c	776	16	0.7	755	9	US-09-990-437-15 Sequence 15, Appl
c	777	16	0.7	755	9	US-09-998-156-15 Sequence 15, Appl
c	778	16	0.7	755	9	US-09-991-157-15 Sequence 15, Appl
c	779	16	0.7	755	9	US-09-991-172-15 Sequence 15, Appl
c	780	16	0.7	755	9	US-09-997-514-15 Sequence 15, Appl
c	781	16	0.7	755	9	US-09-997-573-15 Sequence 15, Appl
c	782	16	0.7	755	9	US-09-990-443-15 Sequence 15, Appl
c	783	16	0.7	755	9	US-09-990-726-15 Sequence 15, Appl
c	784	16	0.7	755	9	US-09-997-559-15 Sequence 15, Appl
c	785	16	0.7	755	9	US-09-997-601-15 Sequence 15, Appl
c	786	16	0.7	755	9	US-09-989-729A-15 Sequence 15, Appl
c	787	16	0.7	755	9	US-09-990-440-15 Sequence 15, Appl
c	788	16	0.7	755	9	US-09-991-854-15 Sequence 15, Appl
c	789	16	0.7	755	9	US-09-997-349-15 Sequence 15, Appl
c	790	16	0.7	755	9	US-09-997-440-15 Sequence 15, Appl
c	791	16	0.7	755	9	US-09-997-628-15 Sequence 15, Appl
c	792	16	0.7	755	9	US-09-997-683-15 Sequence 15, Appl
c	793	16	0.7	755	9	US-09-993-469-15 Sequence 15, Appl
c	794	16	0.7	755	9	US-09-993-748-15 Sequence 15, Appl
c	795	16	0.7	755	9	US-09-997-542-15 Sequence 15, Appl
c	796	16	0.7	755	9	US-09-990-427-15 Sequence 15, Appl
c	797	16	0.7	755	9	US-09-990-439-15 Sequence 15, Appl
c	798	16	0.7	755	9	US-09-989-328-15 Sequence 15, Appl
c	799	16	0.7	755	9	US-09-993-583-15 Sequence 15, Appl
c	800	16	0.7	755	9	US-09-941-992-15 Sequence 15, Appl
c	801	16	0.7	755	9	US-09-992-521-15 Sequence 15, Appl
c	802	16	0.7	755	9	US-09-997-333-15 Sequence 15, Appl
c	803	16	0.7	755	9	US-09-997-384-15 Sequence 15, Appl
c	804	16	0.7	755	10	US-09-989-722-15 Sequence 15, Appl
c	805	16	0.7	755	10	US-09-989-723-15 Sequence 15, Appl
c	806	16	0.7	755	10	US-09-989-279-15 Sequence 15, Appl
c	807	16	0.7	755	10	US-09-989-727-15 Sequence 15, Appl
c	808	16	0.7	755	10	US-09-989-731-15 Sequence 15, Appl
c	809	16	0.7	755	10	US-09-989-732-15 Sequence 15, Appl
c	810	16	0.7	755	10	US-09-991-073-15 Sequence 15, Appl
c	811	16	0.7	755	10	US-09-990-442-15 Sequence 15, Appl
c	812	16	0.7	755	10	US-09-991-163-15 Sequence 15, Appl
c	813	16	0.7	755	10	US-09-993-604-15 Sequence 15, Appl
c	814	16	0.7	755	10	US-09-990-456-15 Sequence 15, Appl
c	815	16	0.7	755	10	US-09-989-721-15 Sequence 15, Appl
c	816	16	0.7	759	10	US-09-770-445-968 Sequence 968, App
	817	16	0.7	762	10	US-09-815-242-4379 Sequence 4379, Ap
	818	16	0.7	765	9	US-10-066-543-119 Sequence 119, App
	819	16	0.7	765	10	US-09-815-242-8176 Sequence 8176, Ap
	820	16	0.7	782	9	US-09-809-391-283 Sequence 283, App
	821	16	0.7	804	10	US-09-770-445-802 Sequence 802, App
	822	16	0.7	833	10	US-09-770-445-686 Sequence 686, App

	823	16	0.7	833	12	US-10-044-090-399	Sequence 399, App
	824	16	0.7	846	9	US-10-001-835-96	Sequence 96, Appl
	825	16	0.7	847	9	US-10-198-846-1684	Sequence 1684, Ap
	826	16	0.7	870	10	US-09-974-300-6506	Sequence 6506, Ap
c	827	16	0.7	906	9	US-10-174-590-541	Sequence 541, App
c	828	16	0.7	906	9	US-10-176-758-541	Sequence 541, App
c	829	16	0.7	906	9	US-10-175-737-541	Sequence 541, App
c	830	16	0.7	906	9	US-10-173-706-541	Sequence 541, App
c	831	16	0.7	906	9	US-10-175-738-541	Sequence 541, App
c	832	16	0.7	906	9	US-10-175-752-541	Sequence 541, App
c	833	16	0.7	906	9	US-10-176-482-541	Sequence 541, App
c	834	16	0.7	906	9	US-10-176-757-541	Sequence 541, App
c	835	16	0.7	906	9	US-10-176-913-541	Sequence 541, App
c	836	16	0.7	906	9	US-10-180-552-541	Sequence 541, App
c	837	16	0.7	906	9	US-10-180-557-541	Sequence 541, App
c	838	16	0.7	906	9	US-10-173-700-541	Sequence 541, App
c	839	16	0.7	906	9	US-10-174-572-541	Sequence 541, App
c	840	16	0.7	906	9	US-10-174-579-541	Sequence 541, App
c	841	16	0.7	906	9	US-10-174-582-541	Sequence 541, App
c	842	16	0.7	906	9	US-10-174-588-541	Sequence 541, App
c	843	16	0.7	906	9	US-10-175-739-541	Sequence 541, App
c	844	16	0.7	906	9	US-10-175-740-541	Sequence 541, App
c	845	16	0.7	906	9	US-10-175-743-541	Sequence 541, App
c	846	16	0.7	906	9	US-10-176-488-541	Sequence 541, App
c	847	16	0.7	906	9	US-10-176-492-541	Sequence 541, App
c	848	16	0.7	906	9	US-10-176-747-541	Sequence 541, App
c	849	16	0.7	906	9	US-10-176-750-541	Sequence 541, App
c	850	16	0.7	906	9	US-10-176-985-541	Sequence 541, App
c	851	16	0.7	906	9	US-10-176-987-541	Sequence 541, App
c	852	16	0.7	906	9	US-10-176-991-541	Sequence 541, App
c	853	16	0.7	906	9	US-10-176-992-541	Sequence 541, App
c	854	16	0.7	906	9	US-10-176-993-541	Sequence 541, App
c	855	16	0.7	906	9	US-10-184-658-541	Sequence 541, App
c	856	16	0.7	906	9	US-10-173-695-541	Sequence 541, App
c	857	16	0.7	906	9	US-10-173-697-541	Sequence 541, App
c	858	16	0.7	906	9	US-10-173-705-541	Sequence 541, App
c	859	16	0.7	906	9	US-10-174-576-541	Sequence 541, App
c	860	16	0.7	906	9	US-10-174-585-541	Sequence 541, App
c	861	16	0.7	906	9	US-10-174-586-541	Sequence 541, App
c	862	16	0.7	906	9	US-10-175-747-541	Sequence 541, App
c	863	16	0.7	906	9	US-10-176-481-541	Sequence 541, App
c	864	16	0.7	906	9	US-10-176-485-541	Sequence 541, App
c	865	16	0.7	906	9	US-10-176-487-541	Sequence 541, App
c	866	16	0.7	906	9	US-10-176-493-541	Sequence 541, App
c	867	16	0.7	906	9	US-10-176-756-541	Sequence 541, App
c	868	16	0.7	906	9	US-10-176-911-541	Sequence 541, App
c	869	16	0.7	906	9	US-10-176-919-541	Sequence 541, App
c	870	16	0.7	906	9	US-10-176-925-541	Sequence 541, App
c	871	16	0.7	906	9	US-10-176-978-541	Sequence 541, App
c	872	16	0.7	906	9	US-10-179-510-541	Sequence 541, App
c	873	16	0.7	906	9	US-10-180-543-541	Sequence 541, App
c	874	16	0.7	906	9	US-10-180-544-541	Sequence 541, App
c	875	16	0.7	906	9	US-10-180-546-541	Sequence 541, App
c	876	16	0.7	906	9	US-10-180-547-541	Sequence 541, App
c	877	16	0.7	906	9	US-10-180-549-541	Sequence 541, App
c	878	16	0.7	906	9	US-10-180-555-541	Sequence 541, App
c	879	16	0.7	906	9	US-10-180-559-541	Sequence 541, App
c	880	16	0.7	906	9	US-10-181-000-541	Sequence 541, App
c	881	16	0.7	906	9	US-10-183-010-541	Sequence 541, App
c	882	16	0.7	906	9	US-10-183-012-541	Sequence 541, App
c	883	16	0.7	906	9	US-10-184-614-541	Sequence 541, App
c	884	16	0.7	906	9	US-10-184-623-541	Sequence 541, App
c	885	16	0.7	906	9	US-10-184-635-541	Sequence 541, App
c	886	16	0.7	906	9	US-10-184-637-541	Sequence 541, App
c	887	16	0.7	906	9	US-10-184-646-541	Sequence 541, App
c	888	16	0.7	906	9	US-10-184-647-541	Sequence 541, App
c	889	16	0.7	906	9	US-10-184-652-541	Sequence 541, App
c	890	16	0.7	906	9	US-10-187-594-541	Sequence 541, App
c	891	16	0.7	906	9	US-10-187-596-541	Sequence 541, App
c	892	16	0.7	906	9	US-10-187-745-541	Sequence 541, App
c	893	16	0.7	906	9	US-10-187-885-541	Sequence 541, App
c	894	16	0.7	906	9	US-10-187-886-541	Sequence 541, App
c	895	16	0.7	906	9	US-10-199-464-541	Sequence 541, App
c	896	16	0.7	906	9	US-10-176-751-541	Sequence 541, App
c	897	16	0.7	906	9	US-10-176-760-541	Sequence 541, App
c	898	16	0.7	906	9	US-10-176-990-541	Sequence 541, App
c	899	16	0.7	906	9	US-10-180-541-541	Sequence 541, App
c	900	16	0.7	906	9	US-10-180-542-541	Sequence 541, App
c	901	16	0.7	906	9	US-10-180-548-541	Sequence 541, App
c	902	16	0.7	906	9	US-10-180-551-541	Sequence 541, App
c	903	16	0.7	906	9	US-10-180-998-541	Sequence 541, App
c	904	16	0.7	906	9	US-10-180-999-541	Sequence 541, App
c	905	16	0.7	906	9	US-10-183-013-541	Sequence 541, App
c	906	16	0.7	906	9	US-10-184-612-541	Sequence 541, App
c	907	16	0.7	906	9	US-10-184-616-541	Sequence 541, App
c	908	16	0.7	906	9	US-10-184-617-541	Sequence 541, App
c	909	16	0.7	906	9	US-10-184-622-541	Sequence 541, App
c	910	16	0.7	906	9	US-10-184-628-541	Sequence 541, App
c	911	16	0.7	906	9	US-10-184-629-541	Sequence 541, App
c	912	16	0.7	906	9	US-10-184-630-541	Sequence 541, App
c	913	16	0.7	906	9	US-10-184-631-541	Sequence 541, App
c	914	16	0.7	906	9	US-10-184-632-541	Sequence 541, App
c	915	16	0.7	906	9	US-10-184-636-541	Sequence 541, App
c	916	16	0.7	906	9	US-10-184-640-541	Sequence 541, App
c	917	16	0.7	906	9	US-10-184-650-541	Sequence 541, App
c	918	16	0.7	906	9	US-10-184-651-541	Sequence 541, App
c	919	16	0.7	906	9	US-10-187-588-541	Sequence 541, App
c	920	16	0.7	906	9	US-10-187-597-541	Sequence 541, App
c	921	16	0.7	906	9	US-10-187-598-541	Sequence 541, App
c	922	16	0.7	906	9	US-10-187-600-541	Sequence 541, App
c	923	16	0.7	906	9	US-10-187-601-541	Sequence 541, App
c	924	16	0.7	906	9	US-10-187-602-541	Sequence 541, App
c	925	16	0.7	906	9	US-10-187-603-541	Sequence 541, App
c	926	16	0.7	906	9	US-10-187-741-541	Sequence 541, App
c	927	16	0.7	906	9	US-10-187-743-541	Sequence 541, App
c	928	16	0.7	906	9	US-10-187-746-541	Sequence 541, App
c	929	16	0.7	906	9	US-10-187-747-541	Sequence 541, App
c	930	16	0.7	906	9	US-10-187-751-541	Sequence 541, App
c	931	16	0.7	906	9	US-10-187-753-541	Sequence 541, App
c	932	16	0.7	906	9	US-10-187-754-541	Sequence 541, App
c	933	16	0.7	906	9	US-10-187-757-541	Sequence 541, App
c	934	16	0.7	906	9	US-10-187-884-541	Sequence 541, App
c	935	16	0.7	906	9	US-10-188-767-541	Sequence 541, App
c	936	16	0.7	906	9	US-10-188-769-541	Sequence 541, App
c	937	16	0.7	906	9	US-10-188-770-541	Sequence 541, App
c	938	16	0.7	906	9	US-10-188-773-541	Sequence 541, App
c	939	16	0.7	906	9	US-10-188-781-541	Sequence 541, App
c	940	16	0.7	906	9	US-10-194-361-541	Sequence 541, App
c	941	16	0.7	906	9	US-10-194-423-541	Sequence 541, App
c	942	16	0.7	906	9	US-10-195-897-541	Sequence 541, App
c	943	16	0.7	906	9	US-10-195-901-541	Sequence 541, App
c	944	16	0.7	906	9	US-10-196-756-541	Sequence 541, App
c	945	16	0.7	906	9	US-10-173-708-541	Sequence 541, App
c	946	16	0.7	906	9	US-10-176-479-541	Sequence 541, App
c	947	16	0.7	906	9	US-10-176-748-541	Sequence 541, App
c	948	16	0.7	906	9	US-10-176-916-541	Sequence 541, App
c	949	16	0.7	906	9	US-10-179-507-541	Sequence 541, App
c	950	16	0.7	906	9	US-10-179-516-541	Sequence 541, App
c	951	16	0.7	906	9	US-10-179-519-541	Sequence 541, App
c	952	16	0.7	906	9	US-10-179-525-541	Sequence 541, App
c	953	16	0.7	906	9	US-10-180-540-541	Sequence 541, App
c	954	16	0.7	906	9	US-10-180-545-541	Sequence 541, App
c	955	16	0.7	906	9	US-10-183-006-541	Sequence 541, App
c	956	16	0.7	906	9	US-10-183-008-541	Sequence 541, App
c	957	16	0.7	906	9	US-10-183-017-541	Sequence 541, App
c	958	16	0.7	906	9	US-10-183-019-541	Sequence 541, App
c	959	16	0.7	906	9	US-10-184-618-541	Sequence 541, App
c	960	16	0.7	906	9	US-10-184-625-541	Sequence 541, App
c	961	16	0.7	906	9	US-10-184-626-541	Sequence 541, App
c	962	16	0.7	906	9	US-10-184-627-541	Sequence 541, App
c	963	16	0.7	906	9	US-10-184-645-541	Sequence 541, App
c	964	16	0.7	906	9	US-10-184-654-541	Sequence 541, App
c	965	16	0.7	906	9	US-10-184-655-541	Sequence 541, App
c	966	16	0.7	906	9	US-10-188-774-541	Sequence 541, App
c	967	16	0.7	906	9	US-10-188-775-541	Sequence 541, App
c	968	16	0.7	906	9	US-10-194-462-541	Sequence 541, App

```
c 969     16     0.7     906   9  US-10-195-902-541          Sequence 541, App
c 970     16     0.7     906   9  US-10-196-743-541          Sequence 541, App
c 971     16     0.7     906   9  US-10-196-745-541          Sequence 541, App
c 972     16     0.7     906   9  US-10-196-760-541          Sequence 541, App
c 973     16     0.7     906   9  US-10-196-762-541          Sequence 541, App
c 974     16     0.7     906   9  US-10-197-695-541          Sequence 541, App
c 975     16     0.7     906   9  US-10-176-484-541          Sequence 541, App
c 976     16     0.7     906   9  US-10-176-753-541          Sequence 541, App
c 977     16     0.7     906   9  US-10-176-917-541          Sequence 541, App
c 978     16     0.7     906   9  US-10-176-982-541          Sequence 541, App
c 979     16     0.7     906   9  US-10-179-506-541          Sequence 541, App
c 980     16     0.7     906   9  US-10-179-513-541          Sequence 541, App
c 981     16     0.7     906   9  US-10-179-514-541          Sequence 541, App
c 982     16     0.7     906   9  US-10-179-522-541          Sequence 541, App
c 983     16     0.7     906   9  US-10-180-556-541          Sequence 541, App
c 984     16     0.7     906   9  US-10-180-560-541          Sequence 541, App
c 985     16     0.7     906   9  US-10-183-015-541          Sequence 541, App
c 986     16     0.7     906   9  US-10-184-615-541          Sequence 541, App
c 987     16     0.7     906   9  US-10-184-620-541          Sequence 541, App
c 988     16     0.7     906   9  US-10-184-643-541          Sequence 541, App
c 989     16     0.7     906   9  US-10-184-656-541          Sequence 541, App
c 990     16     0.7     906   9  US-10-192-010-541          Sequence 541, App
c 991     16     0.7     906   9  US-10-195-894-541          Sequence 541, App
c 992     16     0.7     906   9  US-10-205-908-541          Sequence 541, App
c 993     16     0.7     906   9  US-10-186-855-541          Sequence 541, App
c 994     16     0.7     906   9  US-10-187-599-541          Sequence 541, App
c 995     16     0.7     906   9  US-10-187-750-541          Sequence 541, App
c 996     16     0.7     906   9  US-10-188-780-541          Sequence 541, App
c 997     16     0.7     906   9  US-10-192-015-541          Sequence 541, App
c 998     16     0.7     906   9  US-10-194-394-541          Sequence 541, App
c 999     16     0.7     906   9  US-10-194-425-541          Sequence 541, App
c1000     16     0.7     906   9  US-10-194-485-541          Sequence 541, App

                              ALIGNMENTS

RESULT 1
US-09-833-381-1475
; Sequence 1475, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1475
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-1475

  Query Match             15.7%;  Score 340;  DB 10;  Length 420;
  Best Local Similarity   99.7%;  Pred. No. 1e-177;
  Matches  390;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        807 CAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGAC 866
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         30 CAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAGTTCTGAC 89

Qy        867 TTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATAC 926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 TTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATAC 149

Qy        927 TAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGAT 986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 TAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGAT 209

Qy        987 GGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCA 1046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCA 269

Qy       1047 TATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTAT 1106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 TATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTAT 329

Qy       1107 TATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTC 1166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 TATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTC 389

Qy       1167 TTTTGGGACTTGTGTAAGAGCTGTTGTTATA 1197
              |||||||||||||||||||||||||||||||
Db        390 TTTTGGGACTTGTGTAAGAGCTGTTGTTATA 420

RESULT 2
US-09-764-891-1170
; Sequence 1170, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1170
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-1170

  Query Match             11.7%;  Score 255;  DB 9;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-130;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1883 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAAT 1942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAAT 89

Qy       1943 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 2002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 149

Qy       2003 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAA 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAA 209

Qy       2063 GAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG 269

Qy       2123 GGTTTTCTTCTTCAC 2137
              |||||||||||||||
Db        270 GGTTTTCTTCTTCAC 284

RESULT 3
US-09-764-891-7481
; Sequence 7481, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
```

```
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7481
;   LENGTH: 2194
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-7481

  Query Match             10.1%;  Score 220;  DB 9;  Length 2194;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-111;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              ||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 4
US-09-764-891-7482
; Sequence 7482, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7482
;   LENGTH: 2195
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-7482

  Query Match             10.1%;  Score 220;  DB 9;  Length 2195;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-111;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              ||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 5
US-09-764-891-7483
; Sequence 7483, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7483
;   LENGTH: 2197
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-7483

  Query Match             10.1%;  Score 220;  DB 9;  Length 2197;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-111;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              ||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 6
US-09-815-242-6893
; Sequence 6893, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6893
;   LENGTH: 1842
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1842)
US-09-815-242-6893

  Query Match              1.0%;  Score 21;  DB 10;  Length 1842;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1879 CGAATTGGGCGTACTGGTCGT 1899
              |||||||||||||||||||||
Db        991 CGAATTGGGCGTACTGGTCGT 1011


RESULT 7
US-10-329-960-1
; Sequence 1, Application US/10329960
; Publication No. US20030099277A1
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
;  TITLE OF INVENTION:  Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P1
;  CURRENT APPLICATION NUMBER: US/10/329,960
;  CURRENT FILING DATE:  2003-01-02
;  PRIOR APPLICATION NUMBER: US 09/643,990
;  PRIOR FILING DATE: 2000-08-23
;  PRIOR APPLICATION NUMBER: US 08/487,429
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830121
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4747)..(4747)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9921)..(9921)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10150)..(10150)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29298)..(29298)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36543)..(36543)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36551)..(36551)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36636)..(36636)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (40808)..(40810)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44416)..(44416)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44905)..(44905)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44975)..(44975)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45593)..(45593)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45732)..(45732)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (47036)..(47036)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51334)..(51334)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51602)..(51602)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51786)..(51786)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51805)..(51805)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (55369)..(55369)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65309)..(65309)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65313)..(65313)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (80024)..(80024)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (100091)..(100091)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (102696)..(102696)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (105121)..(105121)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (107248)..(107248)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (117136)..(117136)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (119750)..(119750)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119924)..(119924)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (120038)..(120038)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (121344)..(121344)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122167)..(122167)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122336)..(122336)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131340)..(131340)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131360)..(131360)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (139910)..(139910)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (140398)..(140398)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152530)..(152530)

  Query Match             1.0%;  Score 21;  DB 9;  Length 1830121;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1879 CGAATTGGGCGTACTGGTCGT 1899
              |||||||||||||||||||||
Db     261968 CGAATTGGGCGTACTGGTCGT 261988

RESULT 8
US-09-878-574-10286
; Sequence 10286, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 10286
;   LENGTH: 244
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701103507H1
US-09-878-574-10286

  Query Match             0.9%;  Score 20;  DB 10;  Length 244;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        686 CAGAAGCAGAAGGAGGAGAA 705
              ||||||||||||||||||||
Db          8 CAGAAGCAGAAGGAGGAGAA 27

RESULT 9
US-09-878-574-12453
; Sequence 12453, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 12453
;   LENGTH: 266
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701065791H1
US-09-878-574-12453

  Query Match             0.9%;  Score 20;  DB 10;  Length 266;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        686 CAGAAGCAGAAGGAGGAGAA 705
              ||||||||||||||||||||
Db          3 CAGAAGCAGAAGGAGGAGAA 22

RESULT 10
US-09-864-761-28835/c
```

```
; Sequence 28835, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28835
;   LENGTH: 138
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132775.14
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.98
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.96
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
;   OTHER INFORMATION: EST_HUMAN HIT: H69214.1, EVALUE 2.00e-24
;   OTHER INFORMATION: NT HIT: AL163283.2, EVALUE 9.00e-14
US-09-864-761-28835

  Query Match             0.9%;  Score 19;  DB 10;  Length 138;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        435 CAGAAGAGGTGGAAGAGGT 453
              |||||||||||||||||||
Db         83 CAGAAGAGGTGGAAGAGGT 65


RESULT 11
US-09-918-995-35717
; Sequence 35717, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 35717
;   LENGTH: 418
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(418)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-35717

  Query Match             0.9%;  Score 19;  DB 9;  Length 418;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1829 AACATGTTATCAATTTTGA 1847
              |||||||||||||||||||
Db        249 AACATGTTATCAATTTTGA 267


RESULT 12
US-09-864-761-12255/c
; Sequence 12255, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12255
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132775.14
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.98
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.96
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
US-09-864-761-12255

  Query Match             0.9%;  Score 19;  DB 10;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        435 CAGAAGAGGTGGAAGAGGT 453
              |||||||||||||||||||
Db        320 CAGAAGAGGTGGAAGAGGT 302

RESULT 13
US-09-925-299-60/c
; Sequence 60, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
;   LENGTH: 1472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (129)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (130)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-60

  Query Match             0.9%;  Score 19;  DB 9;  Length 1472;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        316 AGCTCTGGTTTCTGGAGAG 334
              |||||||||||||||||||
Db        354 AGCTCTGGTTTCTGGAGAG 336

RESULT 14
US-09-925-299-60/c
; Sequence 60, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
;   LENGTH: 1472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (129)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (130)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-60

  Query Match             0.9%;  Score 19;  DB 10;  Length 1472;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        316 AGCTCTGGTTTCTGGAGAG 334
              |||||||||||||||||||
Db        354 AGCTCTGGTTTCTGGAGAG 336

RESULT 15
US-09-925-301-185
; Sequence 185, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 185
;   LENGTH: 1987
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (523)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-185

  Query Match             0.9%;  Score 19;  DB 10;  Length 1987;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1434 TATGTTCAGTGCAACTTTT 1452
              |||||||||||||||||||
```

```
Db       1074 TATGTTCAGTGCAACTTTT 1092

Search completed: June 10, 2003, 16:35:47
Job time : 335 secs
```

OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 12:19:45 ; Search time 130 Seconds
(without alignments)
5123.861 Million cell updates/sec

Title: US-09-714-865-15
Perfect score: 2172
Sequence: 1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	0.9	2319	3	US-09-058-489-90	Sequence 90, Appl
	2	19	0.9	2850	4	US-09-184-964-2	Sequence 2, Appli
	3	19	0.9	3408	3	US-09-058-489-14	Sequence 14, Appl
	4	19	0.9	4416	3	US-09-058-489-17	Sequence 17, Appl
	5	19	0.9	5322	3	US-09-058-489-13	Sequence 13, Appl
	6	18	0.8	339	4	US-08-651-155B-116	Sequence 116, App
	7	18	0.8	421	4	US-09-222-575-116	Sequence 116, App
	8	18	0.8	532	4	US-09-318-443-3	Sequence 3, Appli
	9	18	0.8	1536	4	US-09-318-443-5	Sequence 5, Appli
	10	18	0.8	1682	4	US-09-318-443-7	Sequence 7, Appli
c	11	18	0.8	1796	4	US-09-255-392-1	Sequence 1, Appli
	12	18	0.8	1876	4	US-09-082-092-5	Sequence 5, Appli
	13	18	0.8	1914	4	US-09-291-922-19	Sequence 19, Appl
c	14	18	0.8	2005	2	US-08-484-200-1	Sequence 1, Appli
c	15	18	0.8	2005	2	US-08-477-493-1	Sequence 1, Appli
c	16	18	0.8	2005	3	US-08-465-375-1	Sequence 1, Appli
c	17	18	0.8	2005	4	US-09-788-070-1	Sequence 1, Appli
	18	18	0.8	3103	3	US-08-826-246-3	Sequence 3, Appli
	19	18	0.8	3103	3	US-08-944-495-3	Sequence 3, Appli
	20	18	0.8	3103	3	US-09-126-640-2	Sequence 2, Appli
	21	18	0.8	3103	4	US-08-925-588-3	Sequence 3, Appli
	22	18	0.8	3103	4	US-09-288-292A-2	Sequence 2, Appli
	23	18	0.8	3111	3	US-09-487-444-3	Sequence 3, Appli
c	24	18	0.8	11120	4	US-09-201-227A-7	Sequence 7, Appli
c	25	17	0.8	624	4	US-09-161-241-6	Sequence 6, Appli
c	26	17	0.8	780	4	US-09-161-241-5	Sequence 5, Appli
	27	17	0.8	838	3	US-09-062-416-19	Sequence 19, Appl
c	28	17	0.8	1196	1	US-08-167-939A-5	Sequence 5, Appli
c	29	17	0.8	1196	1	US-08-567-538-5	Sequence 5, Appli
	30	17	0.8	1237	1	US-08-197-792-40	Sequence 40, Appl
	31	17	0.8	1237	1	US-08-459-850-40	Sequence 40, Appl
	32	17	0.8	1237	1	US-08-459-214-40	Sequence 40, Appl
	33	17	0.8	1254	5	PCT-US96-05320A-894	Sequence 894, App
	34	17	0.8	1401	2	US-08-811-897A-32	Sequence 32, Appl
	35	17	0.8	1401	2	US-08-855-213-32	Sequence 32, Appl
	36	17	0.8	1401	4	US-09-201-474-32	Sequence 32, Appl
c	37	17	0.8	1446	4	US-09-537-357-8	Sequence 8, Appli
	38	17	0.8	1485	2	US-08-811-897A-33	Sequence 33, Appl
	39	17	0.8	1485	2	US-08-855-213-33	Sequence 33, Appl
	40	17	0.8	1485	4	US-09-201-474-33	Sequence 33, Appl
c	41	17	0.8	1578	1	US-08-188-582-29	Sequence 29, Appl
c	42	17	0.8	1578	1	US-08-646-715-29	Sequence 29, Appl
c	43	17	0.8	1617	4	US-09-532-594B-19	Sequence 19, Appl
c	44	17	0.8	1715	4	US-09-537-357-56	Sequence 56, Appl
c	45	17	0.8	1800	4	US-09-532-594B-17	Sequence 17, Appl
	46	17	0.8	1869	2	US-08-811-897A-40	Sequence 40, Appl
	47	17	0.8	1869	2	US-08-855-213-40	Sequence 40, Appl
	48	17	0.8	1869	4	US-09-201-474-40	Sequence 40, Appl
	49	17	0.8	1924	3	US-08-961-083-159	Sequence 159, App
	50	17	0.8	1929	2	US-08-746-283-4	Sequence 4, Appli
	51	17	0.8	2179	4	US-08-858-207A-37	Sequence 37, Appl
c	52	17	0.8	2208	4	US-09-532-594B-5	Sequence 5, Appli
	53	17	0.8	2386	2	US-08-746-283-2	Sequence 2, Appli
	54	17	0.8	2386	2	US-08-746-257A-2	Sequence 2, Appli
	55	17	0.8	2386	4	US-09-347-878-47	Sequence 47, Appl
	56	17	0.8	2548	2	US-08-811-897A-41	Sequence 41, Appl
	57	17	0.8	2548	2	US-08-855-213-41	Sequence 41, Appl
	58	17	0.8	2548	4	US-09-201-474-41	Sequence 41, Appl
c	59	17	0.8	2608	1	US-08-160-861-1	Sequence 1, Appli
c	60	17	0.8	2846	4	US-08-915-795-1	Sequence 1, Appli
	61	17	0.8	3124	4	US-09-325-932A-131	Sequence 131, App
	62	17	0.8	3393	4	US-09-104-324B-1	Sequence 1, Appli
	63	17	0.8	3393	4	US-09-162-713-1	Sequence 1, Appli
	64	17	0.8	3752	4	US-08-961-527-208	Sequence 208, App
c	65	17	0.8	4767	4	US-09-532-594B-1	Sequence 1, Appli
c	66	17	0.8	6226	1	US-08-542-363-1	Sequence 1, Appli
c	67	17	0.8	6226	4	US-09-100-089-1	Sequence 1, Appli
c	68	17	0.8	6226	4	US-09-670-827-1	Sequence 1, Appli
c	69	17	0.8	8512	4	US-08-961-527-129	Sequence 129, App
	70	17	0.8	42571	4	US-09-810-347-3	Sequence 3, Appli
c	71	17	0.8	246240	2	US-08-724-394A-20	Sequence 20, Appl
c	72	17	0.8	246240	2	US-08-724-394A-21	Sequence 21, Appl
c	73	17	0.8	246240	2	US-08-724-394A-22	Sequence 22, Appl
c	74	17	0.8	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	75	17	0.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	76	16	0.7	160	3	US-08-466-548B-5	Sequence 5, Appli
c	77	16	0.7	160	3	US-07-998-226F-5	Sequence 5, Appli
	78	16	0.7	178	1	US-08-248-474-62	Sequence 62, Appl
	79	16	0.7	178	3	US-08-756-849-62	Sequence 62, Appl
c	80	16	0.7	224	4	US-09-397-787-105	Sequence 105, App
c	81	16	0.7	231	4	US-09-134-001C-1540	Sequence 1540, Ap
	82	16	0.7	252	4	US-08-905-223-232	Sequence 232, App
	83	16	0.7	269	4	US-09-149-476-142	Sequence 142, App
	84	16	0.7	350	4	US-09-118-554-19	Sequence 19, Appl
	85	16	0.7	350	4	US-09-118-627-19	Sequence 19, Appl
	86	16	0.7	350	4	US-09-602-877A-19	Sequence 19, Appl
	87	16	0.7	446	4	US-09-328-111-629	Sequence 629, App
	88	16	0.7	501	4	US-09-339-913B-97	Sequence 97, Appl
	89	16	0.7	501	4	US-09-339-904A-97	Sequence 97, Appl
	90	16	0.7	501	4	US-08-769-062B-97	Sequence 97, Appl
	91	16	0.7	501	4	US-09-344-002B-97	Sequence 97, Appl
	92	16	0.7	501	4	US-09-559-565C-97	Sequence 97, Appl
	93	16	0.7	565	4	US-09-712-016-56	Sequence 56, Appl
	94	16	0.7	641	4	US-08-961-527-359	Sequence 359, App
	95	16	0.7	702	4	US-09-134-001C-806	Sequence 806, App
	96	16	0.7	782	4	US-09-149-476-283	Sequence 283, App
	97	16	0.7	934	2	US-08-867-030B-4	Sequence 4, Appli
	98	16	0.7	934	5	PCT-US95-06119-4	Sequence 4, Appli
c	99	16	0.7	1181	4	US-08-858-207A-124	Sequence 124, App
	100	16	0.7	1191	4	US-09-181-336-18	Sequence 18, Appl

	101	16	0.7	1245	4	US-09-504-358-7	Sequence 7, Appli
	102	16	0.7	1245	4	US-09-954-314-7	Sequence 7, Appli
c	103	16	0.7	1249	3	US-08-985-950-1	Sequence 1, Appli
c	104	16	0.7	1314	4	US-09-134-001C-1971	Sequence 1971, Ap
	105	16	0.7	1332	4	US-09-134-001C-2078	Sequence 2078, Ap
	106	16	0.7	1467	5	PCT-US94-03437-3	Sequence 3, Appli
	107	16	0.7	1500	4	US-09-134-001C-1231	Sequence 1231, Ap
c	108	16	0.7	1519	4	US-09-002-361-4	Sequence 4, Appli
	109	16	0.7	1571	3	US-08-445-463B-3	Sequence 3, Appli
	110	16	0.7	1571	3	US-08-445-464C-3	Sequence 3, Appli
	111	16	0.7	1571	4	US-08-044-857D-3	Sequence 3, Appli
c	112	16	0.7	1666	4	US-09-134-607A-8	Sequence 8, Appli
c	113	16	0.7	1666	4	US-09-134-607A-13	Sequence 13, Appl
c	114	16	0.7	1740	4	US-09-134-607A-10	Sequence 10, Appl
c	115	16	0.7	1740	4	US-09-134-607A-12	Sequence 12, Appl
c	116	16	0.7	1740	4	US-09-134-607A-16	Sequence 16, Appl
c	117	16	0.7	1905	4	US-09-347-878-33	Sequence 33, Appl
c	118	16	0.7	1907	1	US-08-462-884A-2	Sequence 2, Appli
c	119	16	0.7	1908	1	US-08-461-881B-2	Sequence 2, Appli
c	120	16	0.7	1908	2	US-09-123-960-2	Sequence 2, Appli
	121	16	0.7	2067	4	US-09-388-743-21	Sequence 21, Appl
c	122	16	0.7	2068	4	US-09-318-448-15	Sequence 15, Appl
	123	16	0.7	2266	1	US-08-453-472-1	Sequence 1, Appli
	124	16	0.7	2266	1	US-08-453-952-1	Sequence 1, Appli
	125	16	0.7	2266	2	US-08-484-993B-42	Sequence 42, Appl
	126	16	0.7	2266	2	US-08-862-903-1	Sequence 1, Appli
	127	16	0.7	2266	2	US-08-484-158B-42	Sequence 42, Appl
	128	16	0.7	2266	2	US-08-484-596A-42	Sequence 42, Appl
	129	16	0.7	2266	2	US-08-480-150A-42	Sequence 42, Appl
	130	16	0.7	2266	3	US-08-458-731-42	Sequence 42, Appl
	131	16	0.7	2266	3	US-08-149-223A-42	Sequence 42, Appl
c	132	16	0.7	2338	4	US-09-232-200-66	Sequence 66, Appl
c	133	16	0.7	2338	4	US-09-232-197-66	Sequence 66, Appl
c	134	16	0.7	2338	4	US-09-232-201-66	Sequence 66, Appl
	135	16	0.7	2526	4	US-09-202-712-1	Sequence 1, Appli
	136	16	0.7	2559	2	US-08-248-839C-10	Sequence 10, Appl
c	137	16	0.7	2876	4	US-09-134-607A-9	Sequence 9, Appli
c	138	16	0.7	2876	4	US-09-134-607A-14	Sequence 14, Appl
c	139	16	0.7	3023	4	US-09-203-453-4	Sequence 4, Appli
c	140	16	0.7	3066	4	US-09-086-912-1	Sequence 1, Appli
c	141	16	0.7	3066	4	US-09-203-453-1	Sequence 1, Appli
	142	16	0.7	3120	4	US-09-085-761A-45	Sequence 45, Appl
	143	16	0.7	3236	4	US-08-961-527-222	Sequence 222, App
c	144	16	0.7	3259	4	US-09-318-448-23	Sequence 23, Appl
c	145	16	0.7	3348	4	US-09-302-620B-94	Sequence 94, Appl
c	146	16	0.7	3510	4	US-08-961-527-53	Sequence 53, Appl
	147	16	0.7	3684	2	US-08-760-075A-17	Sequence 17, Appl
	148	16	0.7	3684	4	US-09-338-546-17	Sequence 17, Appl
	149	16	0.7	3684	4	US-09-659-084-17	Sequence 17, Appl
c	150	16	0.7	3889	4	US-09-484-970B-39	Sequence 39, Appl
	151	16	0.7	4080	2	US-08-710-249-3	Sequence 3, Appli
	152	16	0.7	4080	4	US-09-220-157A-3	Sequence 3, Appli
	153	16	0.7	4117	4	US-08-961-527-100	Sequence 100, App
c	154	16	0.7	4503	2	US-08-770-301A-2	Sequence 2, Appli
c	155	16	0.7	4503	3	US-09-175-581-2	Sequence 2, Appli
c	156	16	0.7	4951	2	US-08-867-030B-5	Sequence 5, Appli
c	157	16	0.7	4951	5	PCT-US95-06119-5	Sequence 5, Appli
	158	16	0.7	5278	4	US-08-961-527-227	Sequence 227, App
	159	16	0.7	5537	1	US-08-135-511-32	Sequence 32, Appl
	160	16	0.7	5537	1	US-08-483-852-9	Sequence 9, Appli
	161	16	0.7	5537	1	US-08-361-458-4	Sequence 4, Appli
	162	16	0.7	5537	1	US-08-477-953-9	Sequence 9, Appli
	163	16	0.7	5537	1	US-08-187-453-32	Sequence 32, Appl
	164	16	0.7	5537	1	US-08-562-985A-5	Sequence 5, Appli
	165	16	0.7	5537	2	US-08-477-952-9	Sequence 9, Appli
c	166	16	0.7	6203	4	US-09-134-218-3	Sequence 3, Appli
c	167	16	0.7	6680	4	US-09-309-572-6	Sequence 6, Appli
c	168	16	0.7	8298	5	PCT-US93-03076-1	Sequence 1, Appli
	169	16	0.7	8657	4	US-08-961-527-44	Sequence 44, Appl
	170	16	0.7	10350	2	US-09-047-125-1	Sequence 1, Appli
c	171	16	0.7	10350	2	US-09-047-125-2	Sequence 2, Appli
	172	16	0.7	10350	3	US-07-736-335E-1	Sequence 1, Appli
c	173	16	0.7	10350	3	US-07-736-335E-2	Sequence 2, Appli

	174	16	0.7	10629	4	US-09-504-358-15	Sequence 15, Appl
	175	16	0.7	10629	4	US-09-954-314-15	Sequence 15, Appl
	176	16	0.7	11832	2	US-08-416-603-1	Sequence 1, Appli
c	177	16	0.7	12146	4	US-09-277-457-27	Sequence 27, Appl
	178	16	0.7	13206	4	US-08-961-527-33	Sequence 33, Appl
	179	16	0.7	14602	1	US-08-597-236-1	Sequence 1, Appli
	180	16	0.7	14602	1	US-08-746-682A-1	Sequence 1, Appli
	181	16	0.7	18627	4	US-08-961-527-113	Sequence 113, App
	182	16	0.7	22108	3	US-09-053-197A-3	Sequence 3, Appli
	183	16	0.7	22108	4	US-09-085-761A-3	Sequence 3, Appli
c	184	16	0.7	25165	4	US-09-453-702B-39	Sequence 39, Appl
c	185	16	0.7	30549	4	US-09-134-001C-322	Sequence 322, App
	186	16	0.7	35524	3	US-08-923-137-1	Sequence 1, Appli
	187	16	0.7	36519	3	US-08-923-137-2	Sequence 2, Appli
c	188	16	0.7	37948	4	US-09-251-645-11	Sequence 11, Appl
c	189	16	0.7	40328	3	US-08-742-185-102	Sequence 102, App
	190	16	0.7	70000	4	US-09-851-896-3	Sequence 3, Appli
	191	16	0.7	98844	4	US-09-791-211-10	Sequence 10, Appl
c	192	16	0.7	98844	4	US-09-791-211-10	Sequence 10, Appl
c	193	15	0.7	20	2	US-08-973-461A-2	Sequence 2, Appli
c	194	15	0.7	20	3	US-08-648-010-2	Sequence 2, Appli
	195	15	0.7	22	4	US-08-642-274D-78	Sequence 78, Appl
	196	15	0.7	22	4	US-08-952-014C-78	Sequence 78, Appl
c	197	15	0.7	27	4	US-08-584-040-6729	Sequence 6729, Ap
	198	15	0.7	30	3	US-08-738-381-35	Sequence 35, Appl
c	199	15	0.7	33	1	US-08-096-182A-19	Sequence 19, Appl
c	200	15	0.7	33	1	US-08-096-182A-21	Sequence 21, Appl
c	201	15	0.7	33	1	US-08-877-109-19	Sequence 19, Appl
c	202	15	0.7	33	1	US-08-877-109-21	Sequence 21, Appl
c	203	15	0.7	33	3	US-08-798-760-19	Sequence 19, Appl
c	204	15	0.7	33	3	US-08-798-760-21	Sequence 21, Appl
c	205	15	0.7	33	5	PCT-US94-08327-19	Sequence 19, Appl
c	206	15	0.7	33	5	PCT-US94-08327-21	Sequence 21, Appl
	207	15	0.7	34	2	US-08-873-479-23	Sequence 23, Appl
	208	15	0.7	34	2	US-08-972-661A-23	Sequence 23, Appl
c	209	15	0.7	36	2	US-08-873-479-22	Sequence 22, Appl
c	210	15	0.7	36	2	US-08-972-661A-22	Sequence 22, Appl
	211	15	0.7	36	4	US-09-374-962-15	Sequence 15, Appl
	212	15	0.7	36	4	US-09-374-962-16	Sequence 16, Appl
c	213	15	0.7	38	3	US-08-604-757-8	Sequence 8, Appli
c	214	15	0.7	38	4	US-09-612-991-8	Sequence 8, Appli
	215	15	0.7	40	4	US-09-045-583-23	Sequence 23, Appl
	216	15	0.7	40	4	US-09-534-185-23	Sequence 23, Appl
	217	15	0.7	175	3	US-09-103-359-9	Sequence 9, Appli
	218	15	0.7	234	4	US-09-332-319-29	Sequence 29, Appl
c	219	15	0.7	240	1	US-08-222-177A-25	Sequence 25, Appl
	220	15	0.7	241	4	US-09-605-785-412	Sequence 412, App
	221	15	0.7	241	4	US-09-439-313-412	Sequence 412, App
	222	15	0.7	241	4	US-09-352-616A-412	Sequence 412, App
	223	15	0.7	259	3	US-09-103-359-13	Sequence 13, Appl
	224	15	0.7	268	4	US-09-332-319-19	Sequence 19, Appl
c	225	15	0.7	282	4	US-09-072-596-256	Sequence 256, App
	226	15	0.7	299	3	US-09-121-979-2	Sequence 2, Appli
	227	15	0.7	299	4	US-09-332-319-2	Sequence 2, Appli
c	228	15	0.7	306	4	US-08-672-213-10	Sequence 10, Appl
c	229	15	0.7	306	4	US-08-672-213-45	Sequence 45, Appl
c	230	15	0.7	306	4	US-08-973-131-34	Sequence 34, Appl
c	231	15	0.7	318	1	US-08-391-339-10	Sequence 10, Appl
c	232	15	0.7	318	1	US-07-809-457A-13	Sequence 13, Appl
c	233	15	0.7	318	1	US-08-476-008-10	Sequence 10, Appl
c	234	15	0.7	318	1	US-08-306-063-10	Sequence 10, Appl
c	235	15	0.7	318	1	US-08-553-943-13	Sequence 13, Appl
c	236	15	0.7	318	1	US-08-484-274A-10	Sequence 10, Appl
c	237	15	0.7	318	1	US-08-833-485-10	Sequence 10, Appl
c	238	15	0.7	318	4	US-09-137-440-10	Sequence 10, Appl
c	239	15	0.7	318	5	PCT-US91-06148A-10	Sequence 10, Appl
c	240	15	0.7	318	5	PCT-US91-09437-13	Sequence 13, Appl
c	241	15	0.7	335	1	US-08-620-467A-4	Sequence 4, Appli
c	242	15	0.7	335	1	US-08-348-572-4	Sequence 4, Appli
c	243	15	0.7	335	3	US-09-041-090B-4	Sequence 4, Appli
c	244	15	0.7	345	4	US-08-209-603E-4	Sequence 4, Appli
c	245	15	0.7	346	4	US-09-196-293-4	Sequence 4, Appli
c	246	15	0.7	378	1	US-08-118-906-1	Sequence 1, Appli

c	247	15	0.7	378	1	US-08-486-196-1	Sequence 1, Appli
c	248	15	0.7	378	1	US-08-488-135-1	Sequence 1, Appli
c	249	15	0.7	378	2	US-08-474-065-1	Sequence 1, Appli
c	250	15	0.7	393	1	US-07-790-199-1	Sequence 1, Appli
c	251	15	0.7	402	1	US-08-476-008-12	Sequence 12, Appl
c	252	15	0.7	402	1	US-08-306-063-12	Sequence 12, Appl
c	253	15	0.7	402	1	US-08-833-485-12	Sequence 12, Appl
c	254	15	0.7	402	4	US-09-137-440-12	Sequence 12, Appl
c	255	15	0.7	402	5	PCT-US91-06148A-12	Sequence 12, Appl
	256	15	0.7	427	4	US-09-397-787-206	Sequence 206, App
	257	15	0.7	431	4	US-09-328-111-122	Sequence 122, App
c	258	15	0.7	435	4	US-09-397-787-172	Sequence 172, App
	259	15	0.7	448	3	US-09-103-359-12	Sequence 12, Appl
	260	15	0.7	450	4	US-09-397-787-173	Sequence 173, App
	261	15	0.7	451	4	US-09-222-575-113	Sequence 113, App
	262	15	0.7	457	4	US-08-991-789A-271	Sequence 271, App
	263	15	0.7	457	4	US-09-062-451-271	Sequence 271, App
c	264	15	0.7	465	4	US-09-134-001C-798	Sequence 798, App
c	265	15	0.7	480	4	US-09-134-001C-1225	Sequence 1225, Ap
	266	15	0.7	485	4	US-09-385-982-31	Sequence 31, Appl
	267	15	0.7	506	4	US-09-370-838-263	Sequence 263, App
	268	15	0.7	519	4	US-08-580-031A-12	Sequence 12, Appl
	269	15	0.7	522	4	US-09-134-001C-1189	Sequence 1189, Ap
	270	15	0.7	538	4	US-09-643-597-311	Sequence 311, App
	271	15	0.7	560	4	US-09-328-111-18	Sequence 18, Appl
c	272	15	0.7	573	4	US-08-672-213-11	Sequence 11, Appl
	273	15	0.7	573	4	US-09-370-838-268	Sequence 268, App
	274	15	0.7	579	2	US-08-611-757-50	Sequence 50, Appl
	275	15	0.7	579	5	PCT-US95-05980-50	Sequence 50, Appl
c	276	15	0.7	590	4	US-09-643-597-132	Sequence 132, App
	277	15	0.7	591	1	US-08-686-878A-5	Sequence 5, Appli
c	278	15	0.7	652	4	US-09-040-984-49	Sequence 49, Appl
c	279	15	0.7	652	4	US-09-123-912-49	Sequence 49, Appl
c	280	15	0.7	652	4	US-09-643-597-49	Sequence 49, Appl
c	281	15	0.7	676	4	US-09-605-785-355	Sequence 355, App
c	282	15	0.7	676	4	US-09-439-313-355	Sequence 355, App
c	283	15	0.7	676	4	US-09-352-616A-355	Sequence 355, App
	284	15	0.7	679	4	US-08-642-274D-51	Sequence 51, Appl
	285	15	0.7	679	4	US-08-952-014C-51	Sequence 51, Appl
c	286	15	0.7	716	4	US-08-998-416-647	Sequence 647, App
	287	15	0.7	731	1	US-08-451-405A-2	Sequence 2, Appli
c	288	15	0.7	739	4	US-09-182-117-27	Sequence 27, Appl
	289	15	0.7	744	4	US-09-134-001C-2773	Sequence 2773, Ap
	290	15	0.7	750	2	US-08-602-359A-29	Sequence 29, Appl
	291	15	0.7	766	4	US-09-385-982-538	Sequence 538, App
c	292	15	0.7	768	1	US-08-464-590A-1	Sequence 1, Appli
c	293	15	0.7	768	3	US-09-093-585-1	Sequence 1, Appli
c	294	15	0.7	773	4	US-09-020-956-3	Sequence 3, Appli
c	295	15	0.7	773	4	US-09-030-607-3	Sequence 3, Appli
c	296	15	0.7	773	4	US-09-605-785-3	Sequence 3, Appli
c	297	15	0.7	773	4	US-09-439-313-3	Sequence 3, Appli
c	298	15	0.7	773	4	US-09-352-616A-3	Sequence 3, Appli
c	299	15	0.7	773	4	US-09-232-149A-3	Sequence 3, Appli
	300	15	0.7	798	1	US-07-688-352C-41	Sequence 41, Appl
	301	15	0.7	798	2	US-08-474-379C-41	Sequence 41, Appl
	302	15	0.7	798	3	US-09-146-249A-41	Sequence 41, Appl
	303	15	0.7	798	3	US-08-206-188B-41	Sequence 41, Appl
	304	15	0.7	798	5	PCT-US91-02714-39	Sequence 39, Appl
	305	15	0.7	801	4	US-09-134-001C-2125	Sequence 2125, Ap
	306	15	0.7	816	4	US-09-020-956-2	Sequence 2, Appli
	307	15	0.7	816	4	US-09-030-607-2	Sequence 2, Appli
	308	15	0.7	816	4	US-09-605-785-2	Sequence 2, Appli
	309	15	0.7	816	4	US-09-439-313-2	Sequence 2, Appli
	310	15	0.7	816	4	US-09-352-616A-2	Sequence 2, Appli
	311	15	0.7	816	4	US-09-232-149A-2	Sequence 2, Appli
c	312	15	0.7	822	4	US-09-134-001C-804	Sequence 804, App
	313	15	0.7	829	3	US-08-961-083-133	Sequence 133, App
	314	15	0.7	839	1	US-08-884-682-2	Sequence 2, Appli
	315	15	0.7	839	2	US-09-096-082-2	Sequence 2, Appli
	316	15	0.7	841	4	US-08-998-416-318	Sequence 318, App
	317	15	0.7	850	2	US-08-560-398-7	Sequence 7, Appli
	318	15	0.7	853	1	US-08-008-216-16	Sequence 16, Appl
c	319	15	0.7	853	1	US-08-008-216-16	Sequence 16, Appl

	320	15	0.7	853	1	US-08-459-569-16	Sequence 16, Appl
c	321	15	0.7	853	1	US-08-459-569-16	Sequence 16, Appl
	322	15	0.7	853	1	US-08-458-831-16	Sequence 16, Appl
c	323	15	0.7	853	1	US-08-458-831-16	Sequence 16, Appl
c	324	15	0.7	872	4	US-09-484-970B-147	Sequence 147, App
	325	15	0.7	885	4	US-09-641-638-105	Sequence 105, App
	326	15	0.7	895	4	US-09-602-877A-91	Sequence 91, Appl
	327	15	0.7	911	4	US-09-171-209-63	Sequence 63, Appl
	328	15	0.7	924	4	US-09-171-209-43	Sequence 43, Appl
	329	15	0.7	945	4	US-09-134-001C-2568	Sequence 2568, Ap
	330	15	0.7	954	3	US-09-418-641-3	Sequence 3, Appli
c	331	15	0.7	970	1	US-08-690-095-2	Sequence 2, Appli
	332	15	0.7	970	3	US-08-888-077A-28	Sequence 28, Appl
c	333	15	0.7	970	3	US-09-113-789-2	Sequence 2, Appli
c	334	15	0.7	981	4	US-09-134-001C-2097	Sequence 2097, Ap
	335	15	0.7	993	3	US-08-969-644-21	Sequence 21, Appl
	336	15	0.7	993	3	US-08-444-189-21	Sequence 21, Appl
	337	15	0.7	993	4	US-08-468-544-21	Sequence 21, Appl
	338	15	0.7	1002	4	US-09-641-638-591	Sequence 591, App
	339	15	0.7	1005	3	US-09-103-359-4	Sequence 4, Appli
	340	15	0.7	1039	1	US-08-458-084-9	Sequence 9, Appli
	341	15	0.7	1039	1	US-08-205-508-9	Sequence 9, Appli
	342	15	0.7	1039	2	US-08-482-148-6	Sequence 6, Appli
	343	15	0.7	1039	5	PCT-US95-02944-6	Sequence 6, Appli
	344	15	0.7	1039	5	PCT-US95-02945-9	Sequence 9, Appli
c	345	15	0.7	1056	4	US-09-134-001C-2039	Sequence 2039, Ap
c	346	15	0.7	1068	4	US-09-134-001C-1881	Sequence 1881, Ap
c	347	15	0.7	1079	2	US-08-827-190-7	Sequence 7, Appli
c	348	15	0.7	1079	4	US-09-170-187-7	Sequence 7, Appli
	349	15	0.7	1080	4	US-09-149-476-229	Sequence 229, App
c	350	15	0.7	1095	4	US-08-928-383B-3	Sequence 3, Appli
c	351	15	0.7	1097	4	US-08-858-207A-203	Sequence 203, App
	352	15	0.7	1131	4	US-08-975-762-16	Sequence 16, Appl
	353	15	0.7	1131	4	US-08-821-324-16	Sequence 16, Appl
	354	15	0.7	1131	4	US-09-295-028-16	Sequence 16, Appl
	355	15	0.7	1131	4	US-09-106-582-16	Sequence 16, Appl
c	356	15	0.7	1143	4	US-09-276-531-103	Sequence 103, App
	357	15	0.7	1149	4	US-09-134-001C-928	Sequence 928, App
c	358	15	0.7	1150	2	US-08-438-439C-1	Sequence 1, Appli
c	359	15	0.7	1150	3	US-08-705-245-19	Sequence 19, Appl
c	360	15	0.7	1150	5	PCT-US96-06665-1	Sequence 1, Appli
	361	15	0.7	1185	4	US-08-975-762-15	Sequence 15, Appl
	362	15	0.7	1185	4	US-08-821-324-15	Sequence 15, Appl
	363	15	0.7	1185	4	US-09-295-028-15	Sequence 15, Appl
	364	15	0.7	1185	4	US-09-106-582-15	Sequence 15, Appl
c	365	15	0.7	1191	4	US-09-134-001C-1743	Sequence 1743, Ap
	366	15	0.7	1197	2	US-08-973-461A-3	Sequence 3, Appli
	367	15	0.7	1197	3	US-08-648-010-3	Sequence 3, Appli
c	368	15	0.7	1223	3	US-08-155-005A-5	Sequence 5, Appli
c	369	15	0.7	1223	4	US-09-363-783-5	Sequence 5, Appli
	370	15	0.7	1251	3	US-09-007-476-1	Sequence 1, Appli
c	371	15	0.7	1258	1	US-08-335-518-1	Sequence 1, Appli
c	372	15	0.7	1258	4	US-08-988-054-1	Sequence 1, Appli
c	373	15	0.7	1298	4	US-08-795-876-32	Sequence 32, Appl
	374	15	0.7	1300	4	US-08-795-876-31	Sequence 31, Appl
	375	15	0.7	1300	4	US-08-795-876-36	Sequence 36, Appl
c	376	15	0.7	1300	4	US-08-795-876-37	Sequence 37, Appl
	377	15	0.7	1335	4	US-09-134-001C-1509	Sequence 1509, Ap
c	378	15	0.7	1354	4	US-08-456-200B-4	Sequence 4, Appli
	379	15	0.7	1356	4	US-09-134-001C-617	Sequence 617, App
c	380	15	0.7	1374	4	US-09-411-687A-24	Sequence 24, Appl
c	381	15	0.7	1377	4	US-09-221-017B-1113	Sequence 1113, Ap
	382	15	0.7	1383	4	US-09-134-001C-142	Sequence 142, App
	383	15	0.7	1417	1	US-08-713-828-2	Sequence 2, Appli
	384	15	0.7	1417	2	US-08-919-627-2	Sequence 2, Appli
	385	15	0.7	1417	2	US-09-096-245-2	Sequence 2, Appli
c	386	15	0.7	1429	3	US-08-155-005A-7	Sequence 7, Appli
c	387	15	0.7	1429	4	US-09-363-783-7	Sequence 7, Appli
c	388	15	0.7	1434	4	US-09-480-921B-9	Sequence 9, Appli
c	389	15	0.7	1434	4	US-09-480-921B-28	Sequence 28, Appl
	390	15	0.7	1448	1	US-08-585-758A-1	Sequence 1, Appli
	391	15	0.7	1448	1	US-08-977-818-1	Sequence 1, Appli
	392	15	0.7	1448	2	US-08-670-274B-1	Sequence 1, Appli

	393	15	0.7	1448	4	US-09-146-187-1	Sequence 1, Appli
	394	15	0.7	1450	2	US-08-692-787-6	Sequence 6, Appli
	395	15	0.7	1452	1	US-08-149-093A-3	Sequence 3, Appli
	396	15	0.7	1452	1	US-08-911-245-3	Sequence 3, Appli
	397	15	0.7	1452	1	US-08-553-058C-3	Sequence 3, Appli
	398	15	0.7	1452	2	US-08-514-451A-3	Sequence 3, Appli
	399	15	0.7	1452	3	US-09-170-331-3	Sequence 3, Appli
	400	15	0.7	1452	4	US-09-510-473-3	Sequence 3, Appli
	401	15	0.7	1452	4	US-09-048-916B-3	Sequence 3, Appli
	402	15	0.7	1456	4	US-09-056-285A-6	Sequence 6, Appli
	403	15	0.7	1458	4	US-09-480-921B-3	Sequence 3, Appli
	404	15	0.7	1458	4	US-09-417-359A-4	Sequence 4, Appli
	405	15	0.7	1468	4	US-09-290-734-23	Sequence 23, Appl
	406	15	0.7	1468	4	US-09-290-734-25	Sequence 25, Appl
	407	15	0.7	1473	4	US-09-056-285A-9	Sequence 9, Appli
c	408	15	0.7	1484	4	US-09-210-168-2	Sequence 2, Appli
	409	15	0.7	1488	4	US-09-134-001C-1436	Sequence 1436, Ap
c	410	15	0.7	1494	4	US-09-413-611A-5	Sequence 5, Appli
c	411	15	0.7	1500	4	US-09-413-611A-7	Sequence 7, Appli
	412	15	0.7	1512	4	US-09-134-001C-643	Sequence 643, App
c	413	15	0.7	1515	3	US-08-747-221B-16	Sequence 16, Appl
	414	15	0.7	1515	3	US-08-747-221B-17	Sequence 17, Appl
c	415	15	0.7	1515	4	US-09-005-051-16	Sequence 16, Appl
	416	15	0.7	1515	4	US-09-005-051-17	Sequence 17, Appl
c	417	15	0.7	1517	1	US-08-032-869A-1	Sequence 1, Appli
c	418	15	0.7	1517	1	US-08-032-869A-2	Sequence 2, Appli
c	419	15	0.7	1517	1	US-08-472-673-1	Sequence 1, Appli
c	420	15	0.7	1517	1	US-08-472-673-2	Sequence 2, Appli
c	421	15	0.7	1517	1	US-08-475-782-1	Sequence 1, Appli
c	422	15	0.7	1517	1	US-08-475-782-2	Sequence 2, Appli
c	423	15	0.7	1517	2	US-08-472-678-1	Sequence 1, Appli
c	424	15	0.7	1517	2	US-08-472-678-2	Sequence 2, Appli
c	425	15	0.7	1517	2	US-08-484-503-1	Sequence 1, Appli
c	426	15	0.7	1517	2	US-08-484-503-2	Sequence 2, Appli
c	427	15	0.7	1518	4	US-08-944-368A-1	Sequence 1, Appli
c	428	15	0.7	1518	4	US-09-820-764-1	Sequence 1, Appli
c	429	15	0.7	1528	4	US-09-362-473-13	Sequence 13, Appl
c	430	15	0.7	1537	4	US-09-318-448-32	Sequence 32, Appl
	431	15	0.7	1567	3	US-08-889-108-16	Sequence 16, Appl
	432	15	0.7	1567	5	PCT-US94-10358-16	Sequence 16, Appl
c	433	15	0.7	1584	4	US-08-928-383B-1	Sequence 1, Appli
c	434	15	0.7	1590	3	US-08-747-221B-23	Sequence 23, Appl
c	435	15	0.7	1590	4	US-09-005-051-23	Sequence 23, Appl
c	436	15	0.7	1598	4	US-09-661-711A-7	Sequence 7, Appli
	437	15	0.7	1615	3	US-09-040-508-1	Sequence 1, Appli
	438	15	0.7	1615	4	US-09-500-654-1	Sequence 1, Appli
	439	15	0.7	1621	4	US-09-020-956-107	Sequence 107, App
	440	15	0.7	1621	4	US-09-030-607-107	Sequence 107, App
	441	15	0.7	1621	4	US-09-605-785-107	Sequence 107, App
	442	15	0.7	1621	4	US-09-439-313-107	Sequence 107, App
	443	15	0.7	1621	4	US-09-352-616A-107	Sequence 107, App
	444	15	0.7	1621	4	US-09-232-149A-107	Sequence 107, App
	445	15	0.7	1630	2	US-08-805-191-1	Sequence 1, Appli
c	446	15	0.7	1650	3	US-08-747-221B-21	Sequence 21, Appl
	447	15	0.7	1650	3	US-08-747-221B-22	Sequence 22, Appl
c	448	15	0.7	1650	3	US-08-747-221B-60	Sequence 60, Appl
	449	15	0.7	1650	3	US-08-747-221B-61	Sequence 61, Appl
c	450	15	0.7	1650	4	US-09-005-051-21	Sequence 21, Appl
	451	15	0.7	1650	4	US-09-005-051-22	Sequence 22, Appl
c	452	15	0.7	1650	4	US-09-005-051-60	Sequence 60, Appl
	453	15	0.7	1650	4	US-09-005-051-61	Sequence 61, Appl
c	454	15	0.7	1651	3	US-08-155-005A-3	Sequence 3, Appli
c	455	15	0.7	1651	4	US-09-363-783-3	Sequence 3, Appli
c	456	15	0.7	1664	4	US-09-045-583-4	Sequence 4, Appli
c	457	15	0.7	1664	4	US-09-534-185-4	Sequence 4, Appli
	458	15	0.7	1670	4	US-09-071-035-267	Sequence 267, App
c	459	15	0.7	1670	5	PCT-US94-01321-71	Sequence 71, Appl
	460	15	0.7	1677	1	US-08-285-440-12	Sequence 12, Appl
	461	15	0.7	1677	1	US-08-630-349-12	Sequence 12, Appl
c	462	15	0.7	1693	4	US-09-661-711A-1	Sequence 1, Appli
	463	15	0.7	1697	4	US-09-556-870A-1	Sequence 1, Appli
c	464	15	0.7	1699	4	US-09-661-711A-9	Sequence 9, Appli
	465	15	0.7	1702	4	US-08-934-254-26	Sequence 26, Appl

c	466	15	0.7	1712	4	US-08-809-326A-8	Sequence 8, Appli
c	467	15	0.7	1729	3	US-08-155-005A-16	Sequence 16, Appl
c	468	15	0.7	1729	4	US-09-363-783-16	Sequence 16, Appl
c	469	15	0.7	1732	4	US-09-449-335-1	Sequence 1, Appli
c	470	15	0.7	1732	4	US-09-449-335-5	Sequence 5, Appli
	471	15	0.7	1744	3	US-08-961-083-131	Sequence 131, App
	472	15	0.7	1758	4	US-09-370-838-25	Sequence 25, Appl
c	473	15	0.7	1762	1	US-08-105-483-300	Sequence 300, App
c	474	15	0.7	1762	1	US-08-709-209-300	Sequence 300, App
c	475	15	0.7	1762	1	US-08-458-101-300	Sequence 300, App
c	476	15	0.7	1763	4	US-09-312-285-4	Sequence 4, Appli
c	477	15	0.7	1763	4	US-09-312-266-4	Sequence 4, Appli
c	478	15	0.7	1763	4	US-09-312-038-2	Sequence 2, Appli
c	479	15	0.7	1763	4	US-09-728-764-4	Sequence 4, Appli
c	480	15	0.7	1763	4	US-09-312-304B-3	Sequence 3, Appli
c	481	15	0.7	1763	4	US-09-728-792-4	Sequence 4, Appli
c	482	15	0.7	1767	3	US-09-199-859-1	Sequence 1, Appli
c	483	15	0.7	1782	1	US-08-153-799-15	Sequence 15, Appl
c	484	15	0.7	1782	4	US-08-769-746-1	Sequence 1, Appli
c	485	15	0.7	1792	3	US-08-747-221B-18	Sequence 18, Appl
	486	15	0.7	1792	3	US-08-747-221B-20	Sequence 20, Appl
c	487	15	0.7	1792	4	US-09-005-051-18	Sequence 18, Appl
	488	15	0.7	1792	4	US-09-005-051-20	Sequence 20, Appl
	489	15	0.7	1792	4	US-09-491-522-4	Sequence 4, Appli
c	490	15	0.7	1807	1	US-08-118-906-13	Sequence 13, Appl
c	491	15	0.7	1807	1	US-08-486-196-13	Sequence 13, Appl
c	492	15	0.7	1807	1	US-08-488-135-13	Sequence 13, Appl
c	493	15	0.7	1807	2	US-08-474-065-13	Sequence 13, Appl
c	494	15	0.7	1825	4	US-09-184-964-1	Sequence 1, Appli
c	495	15	0.7	1827	4	US-08-897-956A-5	Sequence 5, Appli
c	496	15	0.7	1830	1	US-07-691-079C-3	Sequence 3, Appli
c	497	15	0.7	1830	1	US-08-433-037-3	Sequence 3, Appli
c	498	15	0.7	1830	2	US-08-702-572-1	Sequence 1, Appli
c	499	15	0.7	1836	1	US-07-929-580B-1	Sequence 1, Appli
c	500	15	0.7	1839	1	US-07-929-580B-4	Sequence 4, Appli
	501	15	0.7	1850	3	US-08-724-466B-3	Sequence 3, Appli
	502	15	0.7	1850	4	US-08-882-164D-3	Sequence 3, Appli
c	503	15	0.7	1855	4	US-08-810-009-1	Sequence 1, Appli
c	504	15	0.7	1862	2	US-08-797-689-1	Sequence 1, Appli
	505	15	0.7	1863	2	US-08-455-073A-3	Sequence 3, Appli
	506	15	0.7	1863	4	US-09-311-626B-5	Sequence 5, Appli
c	507	15	0.7	1874	4	US-09-167-299-1	Sequence 1, Appli
	508	15	0.7	1890	4	US-09-360-545-77	Sequence 77, Appl
c	509	15	0.7	1894	5	PCT-US91-08177-10	Sequence 10, Appl
	510	15	0.7	1902	1	US-07-688-352C-43	Sequence 43, Appl
	511	15	0.7	1902	2	US-08-474-379C-43	Sequence 43, Appl
	512	15	0.7	1902	3	US-09-146-249A-43	Sequence 43, Appl
	513	15	0.7	1902	3	US-08-206-188B-43	Sequence 43, Appl
	514	15	0.7	1902	5	PCT-US91-02714-40	Sequence 40, Appl
	515	15	0.7	1939	1	US-08-212-188-3	Sequence 3, Appli
	516	15	0.7	1939	3	US-08-970-725-3	Sequence 3, Appli
	517	15	0.7	1939	5	PCT-US95-02708-3	Sequence 3, Appli
	518	15	0.7	1944	4	US-09-377-557-19	Sequence 19, Appl
c	519	15	0.7	1954	4	US-08-809-326A-13	Sequence 13, Appl
c	520	15	0.7	1982	3	US-08-747-221B-13	Sequence 13, Appl
	521	15	0.7	1982	3	US-08-747-221B-15	Sequence 15, Appl
c	522	15	0.7	1982	4	US-09-005-051-13	Sequence 13, Appl
	523	15	0.7	1982	4	US-09-005-051-15	Sequence 15, Appl
	524	15	0.7	1987	2	US-08-990-379-1	Sequence 1, Appli
c	525	15	0.7	1991	4	US-08-235-836C-69	Sequence 69, Appl
c	526	15	0.7	1991	4	US-08-235-836C-77	Sequence 77, Appl
	527	15	0.7	1993	2	US-08-990-379-2	Sequence 2, Appli
c	528	15	0.7	2000	1	US-09-041-075A-4	Sequence 4, Appli
c	529	15	0.7	2000	1	US-09-041-075A-6	Sequence 6, Appli
	530	15	0.7	2032	4	US-09-500-063-1	Sequence 1, Appli
c	531	15	0.7	2040	3	US-08-604-789B-1	Sequence 1, Appli
c	532	15	0.7	2040	3	US-08-604-789B-12	Sequence 12, Appl
c	533	15	0.7	2040	4	US-09-312-721A-1	Sequence 1, Appli
c	534	15	0.7	2040	4	US-09-312-721A-12	Sequence 12, Appl
c	535	15	0.7	2064	1	US-08-164-839-30	Sequence 30, Appl
c	536	15	0.7	2064	1	US-08-164-839-32	Sequence 32, Appl
c	537	15	0.7	2064	1	US-08-583-799-30	Sequence 30, Appl
c	538	15	0.7	2064	1	US-08-583-799-32	Sequence 32, Appl

c	539	15	0.7	2104	1	US-08-592-126-96	Sequence 96, Appl
c	540	15	0.7	2107	4	US-08-235-836C-73	Sequence 73, Appl
c	541	15	0.7	2112	4	US-09-232-160-12	Sequence 12, Appl
c	542	15	0.7	2129	4	US-08-975-762-39	Sequence 39, Appl
c	543	15	0.7	2129	4	US-09-295-028-39	Sequence 39, Appl
c	544	15	0.7	2129	4	US-09-106-582-39	Sequence 39, Appl
c	545	15	0.7	2144	3	US-08-747-221B-57	Sequence 57, Appl
	546	15	0.7	2144	3	US-08-747-221B-59	Sequence 59, Appl
c	547	15	0.7	2144	4	US-09-005-051-57	Sequence 57, Appl
	548	15	0.7	2144	4	US-09-005-051-59	Sequence 59, Appl
c	549	15	0.7	2148	1	US-08-164-839-69	Sequence 69, Appl
c	550	15	0.7	2148	1	US-08-164-839-71	Sequence 71, Appl
c	551	15	0.7	2148	1	US-08-583-799-69	Sequence 69, Appl
c	552	15	0.7	2148	1	US-08-583-799-71	Sequence 71, Appl
	553	15	0.7	2170	4	US-09-480-921B-14	Sequence 14, Appl
	554	15	0.7	2178	2	US-08-474-379C-80	Sequence 80, Appl
	555	15	0.7	2178	3	US-09-146-249A-80	Sequence 80, Appl
	556	15	0.7	2178	3	US-08-206-188B-80	Sequence 80, Appl
	557	15	0.7	2179	1	US-08-487-886-1	Sequence 1, Appli
	558	15	0.7	2179	2	US-08-531-070A-1	Sequence 1, Appli
	559	15	0.7	2179	3	US-08-482-855-1	Sequence 1, Appli
	560	15	0.7	2179	4	US-08-474-986-1	Sequence 1, Appli
	561	15	0.7	2181	4	US-09-299-041-1	Sequence 1, Appli
c	562	15	0.7	2192	1	US-08-399-696-125	Sequence 125, App
	563	15	0.7	2193	4	US-09-427-261-2	Sequence 2, Appli
	564	15	0.7	2193	4	US-09-427-261-3	Sequence 3, Appli
c	565	15	0.7	2196	1	US-08-090-523-29	Sequence 29, Appl
c	566	15	0.7	2196	1	US-08-334-639-1	Sequence 1, Appli
c	567	15	0.7	2196	1	US-08-398-627-29	Sequence 29, Appl
c	568	15	0.7	2196	1	US-08-406-858-24	Sequence 24, Appl
c	569	15	0.7	2196	5	PCT-US94-05275-24	Sequence 24, Appl
	570	15	0.7	2214	2	US-08-484-993B-1	Sequence 1, Appli
	571	15	0.7	2214	2	US-08-484-158B-1	Sequence 1, Appli
	572	15	0.7	2214	2	US-08-484-596A-1	Sequence 1, Appli
	573	15	0.7	2214	2	US-08-480-150A-1	Sequence 1, Appli
	574	15	0.7	2214	3	US-08-458-731-1	Sequence 1, Appli
	575	15	0.7	2214	3	US-08-149-223A-1	Sequence 1, Appli
c	576	15	0.7	2214	4	US-08-943-731-57	Sequence 57, Appl
	577	15	0.7	2215	1	US-08-285-440-14	Sequence 14, Appl
	578	15	0.7	2215	1	US-08-630-349-14	Sequence 14, Appl
	579	15	0.7	2235	4	US-09-484-970B-28	Sequence 28, Appl
c	580	15	0.7	2247	1	US-08-046-508-1	Sequence 1, Appli
c	581	15	0.7	2248	1	US-08-369-780-1	Sequence 1, Appli
c	582	15	0.7	2248	1	US-08-475-682-1	Sequence 1, Appli
c	583	15	0.7	2248	1	US-08-780-833-1	Sequence 1, Appli
c	584	15	0.7	2248	1	US-08-636-036-1	Sequence 1, Appli
c	585	15	0.7	2248	3	US-08-918-509-1	Sequence 1, Appli
c	586	15	0.7	2248	3	US-09-108-262-1	Sequence 1, Appli
	587	15	0.7	2296	2	US-08-899-336-2	Sequence 2, Appli
	588	15	0.7	2311	4	US-09-624-693A-14	Sequence 14, Appl
c	589	15	0.7	2313	4	US-09-370-838-157	Sequence 157, App
	590	15	0.7	2344	1	US-07-695-472B-35	Sequence 35, Appl
	591	15	0.7	2347	1	US-08-250-740-32	Sequence 32, Appl
	592	15	0.7	2347	1	US-07-695-472B-1	Sequence 1, Appli
	593	15	0.7	2347	1	US-07-695-472B-34	Sequence 34, Appl
	594	15	0.7	2347	1	US-07-695-472B-36	Sequence 36, Appl
	595	15	0.7	2352	2	US-08-922-837-1	Sequence 1, Appli
	596	15	0.7	2352	4	US-09-351-550-1	Sequence 1, Appli
c	597	15	0.7	2376	2	US-08-760-745-4	Sequence 4, Appli
	598	15	0.7	2381	2	US-08-484-993B-9	Sequence 9, Appli
	599	15	0.7	2381	2	US-08-484-158B-9	Sequence 9, Appli
	600	15	0.7	2381	2	US-08-484-596A-9	Sequence 9, Appli
	601	15	0.7	2381	2	US-08-480-150A-9	Sequence 9, Appli
	602	15	0.7	2381	3	US-08-458-731-9	Sequence 9, Appli
	603	15	0.7	2381	3	US-08-149-223A-9	Sequence 9, Appli
c	604	15	0.7	2382	1	US-08-256-938-1	Sequence 1, Appli
c	605	15	0.7	2416	4	US-09-513-783A-177	Sequence 177, App
c	606	15	0.7	2434	4	US-09-272-496-1	Sequence 1, Appli
c	607	15	0.7	2439	4	US-09-661-711A-25	Sequence 25, Appl
c	608	15	0.7	2444	3	US-08-920-610-1	Sequence 1, Appli
c	609	15	0.7	2455	1	US-08-256-938-3	Sequence 3, Appli
c	610	15	0.7	2455	2	US-08-797-689-15	Sequence 15, Appl
c	611	15	0.7	2462	4	US-09-661-711A-3	Sequence 3, Appli

	612	15	0.7	2499	4	US-09-276-531-133	Sequence 133, App
	613	15	0.7	2505	2	US-08-890-094-1	Sequence 1, Appli
	614	15	0.7	2547	3	US-08-508-761B-1	Sequence 1, Appli
c	615	15	0.7	2577	4	US-09-605-785-552	Sequence 552, App
c	616	15	0.7	2601	3	US-09-039-773A-2	Sequence 2, Appli
	617	15	0.7	2606	4	US-09-252-816A-7	Sequence 7, Appli
	618	15	0.7	2628	2	US-08-696-944-1	Sequence 1, Appli
c	619	15	0.7	2640	4	US-09-312-285-3	Sequence 3, Appli
c	620	15	0.7	2640	4	US-09-312-266-3	Sequence 3, Appli
c	621	15	0.7	2640	4	US-09-312-038-1	Sequence 1, Appli
c	622	15	0.7	2640	4	US-09-728-764-3	Sequence 3, Appli
c	623	15	0.7	2640	4	US-09-312-304B-1	Sequence 1, Appli
c	624	15	0.7	2640	4	US-09-728-792-3	Sequence 3, Appli
	625	15	0.7	2706	2	US-08-454-549-1	Sequence 1, Appli
	626	15	0.7	2706	3	US-08-454-552-1	Sequence 1, Appli
	627	15	0.7	2706	3	US-08-676-351-1	Sequence 1, Appli
	628	15	0.7	2712	1	US-08-346-455B-37	Sequence 37, Appl
	629	15	0.7	2712	3	US-08-977-221-37	Sequence 37, Appl
	630	15	0.7	2712	4	US-09-483-831B-37	Sequence 37, Appl
	631	15	0.7	2712	5	PCT-US95-06613-37	Sequence 37, Appl
c	632	15	0.7	2728	4	US-09-232-160-1	Sequence 1, Appli
c	633	15	0.7	2773	4	US-09-643-597-358	Sequence 358, App
c	634	15	0.7	2784	4	US-09-643-597-168	Sequence 168, App
c	635	15	0.7	2875	3	US-09-255-893-1	Sequence 1, Appli
	636	15	0.7	2935	4	US-09-480-921B-27	Sequence 27, Appl
c	637	15	0.7	2955	4	US-08-897-956A-4	Sequence 4, Appli
c	638	15	0.7	2970	4	US-09-193-562D-31	Sequence 31, Appl
	639	15	0.7	2995	1	US-08-592-126-85	Sequence 85, Appl
	640	15	0.7	3039	2	US-08-680-326-27	Sequence 27, Appl
c	641	15	0.7	3095	6	5231168-1	Patent No. 5231168
	642	15	0.7	3104	1	US-08-346-455B-66	Sequence 66, Appl
	643	15	0.7	3104	3	US-08-977-221-66	Sequence 66, Appl
	644	15	0.7	3104	4	US-09-483-831B-66	Sequence 66, Appl
	645	15	0.7	3104	5	PCT-US95-06613-66	Sequence 66, Appl
c	646	15	0.7	3123	4	US-09-134-001C-2528	Sequence 2528, Ap
	647	15	0.7	3130	2	US-08-474-379C-62	Sequence 62, Appl
	648	15	0.7	3130	3	US-09-146-249A-62	Sequence 62, Appl
	649	15	0.7	3130	3	US-08-206-188B-62	Sequence 62, Appl
	650	15	0.7	3140	4	US-08-961-527-158	Sequence 158, App
c	651	15	0.7	3143	1	US-08-485-621-1	Sequence 1, Appli
c	652	15	0.7	3143	2	US-08-973-831-1	Sequence 1, Appli
c	653	15	0.7	3143	5	PCT-US96-09530A-1	Sequence 1, Appli
	654	15	0.7	3153	4	US-09-175-928-9	Sequence 9, Appli
	655	15	0.7	3182	4	US-09-221-017B-1029	Sequence 1029, Ap
	656	15	0.7	3251	1	US-08-346-455B-68	Sequence 68, Appl
	657	15	0.7	3251	3	US-08-977-221-68	Sequence 68, Appl
	658	15	0.7	3251	4	US-09-483-831B-68	Sequence 68, Appl
	659	15	0.7	3251	5	PCT-US95-06613-68	Sequence 68, Appl
	660	15	0.7	3353	1	US-08-453-695A-109	Sequence 109, App
	661	15	0.7	3353	1	US-08-268-161A-109	Sequence 109, App
	662	15	0.7	3353	2	US-08-453-702A-109	Sequence 109, App
	663	15	0.7	3353	4	US-09-099-639-109	Sequence 109, App
	664	15	0.7	3353	5	PCT-US95-08071-109	Sequence 109, App
c	665	15	0.7	3362	4	US-09-643-597-167	Sequence 167, App
	666	15	0.7	3411	4	US-09-146-969-4	Sequence 4, Appli
	667	15	0.7	3439	4	US-09-252-816A-6	Sequence 6, Appli
	668	15	0.7	3441	4	US-09-026-033-17	Sequence 17, Appl
c	669	15	0.7	3469	4	US-09-221-017B-106	Sequence 106, App
	670	15	0.7	3479	3	US-08-714-918-7	Sequence 7, Appli
	671	15	0.7	3479	4	US-09-265-315-7	Sequence 7, Appli
	672	15	0.7	3479	4	US-09-265-315-7	Sequence 7, Appli
	673	15	0.7	3479	4	US-09-266-417-7	Sequence 7, Appli
c	674	15	0.7	3497	1	US-08-530-950-5	Sequence 5, Appli
c	675	15	0.7	3497	4	US-09-149-879-5	Sequence 5, Appli
c	676	15	0.7	3498	3	US-08-888-429A-5	Sequence 5, Appli
c	677	15	0.7	3553	1	US-08-530-950-7	Sequence 7, Appli
c	678	15	0.7	3553	4	US-09-149-879-7	Sequence 7, Appli
c	679	15	0.7	3554	3	US-08-888-429A-7	Sequence 7, Appli
c	680	15	0.7	3563	4	US-09-041-886-20	Sequence 20, Appl
c	681	15	0.7	3576	1	US-08-530-950-9	Sequence 9, Appli
c	682	15	0.7	3576	3	US-09-358-382-1	Sequence 1, Appli
c	683	15	0.7	3576	3	US-08-888-429A-9	Sequence 9, Appli
c	684	15	0.7	3576	4	US-09-149-879-9	Sequence 9, Appli

c	685	15	0.7	3587	2	US-08-874-186-91	Sequence 91, Appl
c	686	15	0.7	3596	2	US-08-779-801-5	Sequence 5, Appli
c	687	15	0.7	3596	4	US-09-298-441-5	Sequence 5, Appli
c	688	15	0.7	3632	2	US-08-779-801-3	Sequence 3, Appli
c	689	15	0.7	3632	2	US-08-779-801-4	Sequence 4, Appli
c	690	15	0.7	3632	4	US-09-298-441-3	Sequence 3, Appli
c	691	15	0.7	3632	4	US-09-298-441-4	Sequence 4, Appli
c	692	15	0.7	3796	1	US-08-920-812-19	Sequence 19, Appl
c	693	15	0.7	3796	1	US-08-920-827-19	Sequence 19, Appl
c	694	15	0.7	3796	1	US-08-921-177-19	Sequence 19, Appl
c	695	15	0.7	3796	1	US-08-362-577C-19	Sequence 19, Appl
c	696	15	0.7	3796	2	US-08-920-828-19	Sequence 19, Appl
	697	15	0.7	3804	3	US-09-356-952-14	Sequence 14, Appl
	698	15	0.7	3871	4	US-09-484-970B-70	Sequence 70, Appl
	699	15	0.7	3935	4	US-09-060-482-1	Sequence 1, Appli
c	700	15	0.7	3951	4	US-09-643-597-160	Sequence 160, App
c	701	15	0.7	3991	4	US-08-506-296B-3	Sequence 3, Appli
	702	15	0.7	3998	4	US-09-066-046-5	Sequence 5, Appli
	703	15	0.7	4004	4	US-09-293-505-8	Sequence 8, Appli
	704	15	0.7	4030	4	US-09-293-505-1	Sequence 1, Appli
c	705	15	0.7	4160	4	US-09-341-678-5	Sequence 5, Appli
	706	15	0.7	4162	2	US-08-459-448A-26	Sequence 26, Appl
	707	15	0.7	4162	3	US-08-459-595A-26	Sequence 26, Appl
	708	15	0.7	4162	3	US-08-459-504B-26	Sequence 26, Appl
	709	15	0.7	4162	3	US-08-459-444-26	Sequence 26, Appl
	710	15	0.7	4162	4	US-09-547-422-26	Sequence 26, Appl
	711	15	0.7	4164	1	US-08-188-582-3	Sequence 3, Appli
	712	15	0.7	4164	1	US-08-646-715-3	Sequence 3, Appli
	713	15	0.7	4165	1	US-07-951-715A-26	Sequence 26, Appl
	714	15	0.7	4174	4	US-09-221-017B-926	Sequence 926, App
	715	15	0.7	4208	4	US-09-214-278-6	Sequence 6, Appli
	716	15	0.7	4208	4	US-09-068-740A-10	Sequence 10, Appl
	717	15	0.7	4391	4	US-09-207-857-1	Sequence 1, Appli
c	718	15	0.7	4434	3	US-08-815-809-1	Sequence 1, Appli
c	719	15	0.7	4440	1	US-08-200-016-4	Sequence 4, Appli
	720	15	0.7	4440	3	US-07-792-600-1	Sequence 1, Appli
	721	15	0.7	4440	3	US-09-157-021-1	Sequence 1, Appli
	722	15	0.7	4440	3	US-09-156-842-1	Sequence 1, Appli
	723	15	0.7	4467	1	US-08-565-907A-1	Sequence 1, Appli
c	724	15	0.7	4467	1	US-08-565-907A-1	Sequence 1, Appli
	725	15	0.7	4467	2	US-08-910-551B-1	Sequence 1, Appli
c	726	15	0.7	4467	2	US-08-910-551B-1	Sequence 1, Appli
	727	15	0.7	4467	2	US-08-909-425A-1	Sequence 1, Appli
c	728	15	0.7	4467	2	US-08-909-425A-1	Sequence 1, Appli
c	729	15	0.7	4472	2	US-08-816-155B-3	Sequence 3, Appli
c	730	15	0.7	4472	3	US-09-079-587-3	Sequence 3, Appli
c	731	15	0.7	4517	5	PCT-US93-06251-83	Sequence 83, Appl
	732	15	0.7	4539	1	US-08-119-512-1	Sequence 1, Appli
	733	15	0.7	4539	1	US-08-488-015B-1	Sequence 1, Appli
	734	15	0.7	4542	3	US-08-814-412-11	Sequence 11, Appl
	735	15	0.7	4545	4	US-09-221-017B-1011	Sequence 1011, Ap
c	736	15	0.7	4615	1	US-08-188-582-1	Sequence 1, Appli
c	737	15	0.7	4615	1	US-08-646-715-1	Sequence 1, Appli
	738	15	0.7	4724	4	US-09-066-046-3	Sequence 3, Appli
	739	15	0.7	4741	1	US-07-695-472B-4	Sequence 4, Appli
	740	15	0.7	4742	1	US-08-250-740-35	Sequence 35, Appl
	741	15	0.7	4757	1	US-07-865-662F-12	Sequence 12, Appl
	742	15	0.7	4757	4	US-08-374-219B-12	Sequence 12, Appl
c	743	15	0.7	4792	4	US-08-961-527-156	Sequence 156, App
	744	15	0.7	4977	4	US-09-071-035-257	Sequence 257, App
	745	15	0.7	4977	4	US-09-071-035-261	Sequence 261, App
	746	15	0.7	4977	4	US-09-071-035-265	Sequence 265, App
c	747	15	0.7	5014	4	US-09-381-862-3	Sequence 3, Appli
	748	15	0.7	5057	3	US-08-651-999A-6	Sequence 6, Appli
	749	15	0.7	5057	4	US-09-385-752-6	Sequence 6, Appli
	750	15	0.7	5063	1	US-08-185-432-1	Sequence 1, Appli
	751	15	0.7	5174	4	US-09-066-047-9	Sequence 9, Appli
c	752	15	0.7	5186	2	US-08-596-366-1	Sequence 1, Appli
c	753	15	0.7	5186	2	US-08-967-104-1	Sequence 1, Appli
c	754	15	0.7	5224	2	US-08-874-186-46	Sequence 46, Appl
	755	15	0.7	5375	3	US-08-757-223-7	Sequence 7, Appli
c	756	15	0.7	5402	4	US-09-221-017B-194	Sequence 194, App
	757	15	0.7	5433	3	US-09-157-021-35	Sequence 35, Appl
	758	15	0.7	5433	3	US-09-156-842-35	Sequence 35, Appl
	759	15	0.7	5458	4	US-09-199-865-2	Sequence 2, Appli
	760	15	0.7	5461	3	US-07-998-289B-7	Sequence 7, Appli
	761	15	0.7	5474	4	US-09-040-738-1	Sequence 1, Appli
	762	15	0.7	5474	4	US-08-652-426A-1	Sequence 1, Appli
	763	15	0.7	5499	1	US-07-695-564-4	Sequence 4, Appli
	764	15	0.7	5499	1	US-08-241-387-4	Sequence 4, Appli
	765	15	0.7	5590	3	US-08-882-046-1	Sequence 1, Appli
	766	15	0.7	5629	1	US-07-695-564-2	Sequence 2, Appli
	767	15	0.7	5629	1	US-08-241-387-2	Sequence 2, Appli
c	768	15	0.7	5688	4	US-09-293-170-7	Sequence 7, Appli
c	769	15	0.7	5761	4	US-09-323-472A-1	Sequence 1, Appli
c	770	15	0.7	5761	4	US-09-323-472A-3	Sequence 3, Appli
c	771	15	0.7	5761	4	US-09-323-472A-11	Sequence 11, Appl
c	772	15	0.7	5762	4	US-09-323-472A-13	Sequence 13, Appl
c	773	15	0.7	5816	4	US-08-857-076-11	Sequence 11, Appl
	774	15	0.7	5837	1	US-07-686-340-1	Sequence 1, Appli
	775	15	0.7	5837	1	US-08-004-139B-1	Sequence 1, Appli
	776	15	0.7	5837	1	US-08-117-491-1	Sequence 1, Appli
	777	15	0.7	5837	1	US-08-271-364A-1	Sequence 1, Appli
	778	15	0.7	5837	2	US-08-811-492-1	Sequence 1, Appli
	779	15	0.7	5837	2	US-08-222-715B-1	Sequence 1, Appli
	780	15	0.7	5837	5	PCT-US96-10545A-1	Sequence 1, Appli
c	781	15	0.7	5856	4	US-09-293-170-1	Sequence 1, Appli
	782	15	0.7	5894	3	US-08-665-259-24	Sequence 24, Appl
	783	15	0.7	5894	3	US-08-762-500-24	Sequence 24, Appl
	784	15	0.7	5912	2	US-08-629-001A-1	Sequence 1, Appli
	785	15	0.7	5912	4	US-08-642-274D-1	Sequence 1, Appli
	786	15	0.7	5912	4	US-08-952-127-1	Sequence 1, Appli
	787	15	0.7	5912	4	US-08-952-014C-1	Sequence 1, Appli
	788	15	0.7	6226	1	US-08-542-363-1	Sequence 1, Appli
	789	15	0.7	6226	4	US-09-100-089-1	Sequence 1, Appli
	790	15	0.7	6226	4	US-09-670-827-1	Sequence 1, Appli
c	791	15	0.7	6273	4	US-08-961-527-21	Sequence 21, Appl
	792	15	0.7	6464	2	US-08-400-159-5	Sequence 5, Appli
	793	15	0.7	6464	3	US-08-611-729A-5	Sequence 5, Appli
	794	15	0.7	6513	1	US-08-338-702-7	Sequence 7, Appli
	795	15	0.7	6513	1	US-08-337-339-7	Sequence 7, Appli
	796	15	0.7	6513	1	US-08-724-095-7	Sequence 7, Appli
	797	15	0.7	6513	5	PCT-US95-14262-7	Sequence 7, Appli
	798	15	0.7	6513	5	PCT-US95-14378-7	Sequence 7, Appli
	799	15	0.7	6519	2	US-08-808-793-24	Sequence 24, Appl
	800	15	0.7	6525	1	US-08-493-092-3	Sequence 3, Appli
	801	15	0.7	6525	1	US-08-508-836A-3	Sequence 3, Appli
	802	15	0.7	6525	3	US-08-762-500-74	Sequence 74, Appl
c	803	15	0.7	6587	4	US-09-293-170-8	Sequence 8, Appli
	804	15	0.7	6692	4	US-09-491-522-1	Sequence 1, Appli
	805	15	0.7	6709	4	US-09-285-601-3	Sequence 3, Appli
c	806	15	0.7	6827	4	US-08-961-527-60	Sequence 60, Appl
c	807	15	0.7	6924	1	US-08-015-973-2	Sequence 2, Appli
c	808	15	0.7	6924	2	US-08-448-164-2	Sequence 2, Appli
c	809	15	0.7	6942	2	US-08-460-309-3	Sequence 3, Appli
c	810	15	0.7	6942	2	US-08-125-077-3	Sequence 3, Appli
c	811	15	0.7	7001	4	US-08-961-527-88	Sequence 88, Appl
c	812	15	0.7	7323	5	PCT-US91-08177-1	Sequence 1, Appli
	813	15	0.7	7350	1	US-07-865-662F-14	Sequence 14, Appl
	814	15	0.7	7350	4	US-08-374-219B-16	Sequence 16, Appl
	815	15	0.7	7408	1	US-08-441-822-1	Sequence 1, Appli
	816	15	0.7	7410	1	US-08-493-092-1	Sequence 1, Appli
	817	15	0.7	7410	1	US-08-508-836A-1	Sequence 1, Appli
c	818	15	0.7	7502	3	US-08-969-644-6	Sequence 6, Appli
c	819	15	0.7	7502	3	US-08-444-189-6	Sequence 6, Appli
c	820	15	0.7	7502	4	US-08-468-544-6	Sequence 6, Appli
	821	15	0.7	7521	4	US-09-004-838-116	Sequence 116, App
c	822	15	0.7	7577	4	US-08-961-527-46	Sequence 46, Appl
c	823	15	0.7	7685	4	US-09-221-017B-1092	Sequence 1092, Ap
c	824	15	0.7	7791	2	US-08-149-097D-23	Sequence 23, Appl
c	825	15	0.7	7791	3	US-08-949-386-23	Sequence 23, Appl
c	826	15	0.7	7791	3	US-08-450-562-23	Sequence 23, Appl
c	827	15	0.7	7791	4	US-08-984-709A-23	Sequence 23, Appl
c	828	15	0.7	7791	4	US-08-450-272-23	Sequence 23, Appl
c	829	15	0.7	7808	2	US-08-149-097D-22	Sequence 22, Appl
c	830	15	0.7	7808	3	US-08-949-386-22	Sequence 22, Appl

c	831	15	0.7	7808	3	US-08-450-562-22	Sequence 22, Appl
c	832	15	0.7	7808	4	US-08-984-709A-22	Sequence 22, Appl
c	833	15	0.7	7808	4	US-08-450-272-22	Sequence 22, Appl
c	834	15	0.7	7904	1	US-08-316-239B-1	Sequence 1, Appli
c	835	15	0.7	7904	1	US-08-316-239B-2	Sequence 2, Appli
c	836	15	0.7	7904	1	US-08-410-005-1	Sequence 1, Appli
c	837	15	0.7	7904	3	US-08-929-140-1	Sequence 1, Appli
c	838	15	0.7	7904	4	US-09-560-579A-1	Sequence 1, Appli
c	839	15	0.7	7941	4	US-09-816-703A-1	Sequence 1, Appli
c	840	15	0.7	8012	4	US-09-182-117-1	Sequence 1, Appli
c	841	15	0.7	8031	4	US-09-643-597-254	Sequence 254, App
	842	15	0.7	8050	4	US-09-491-362-11	Sequence 11, Appl
	843	15	0.7	8050	4	US-09-874-562-11	Sequence 11, Appl
c	844	15	0.7	8220	2	US-08-568-459A-11	Sequence 11, Appl
c	845	15	0.7	8220	2	US-08-487-826B-11	Sequence 11, Appl
c	846	15	0.7	8220	4	US-09-210-288-11	Sequence 11, Appl
c	847	15	0.7	8418	4	US-09-182-117-5	Sequence 5, Appli
c	848	15	0.7	8651	4	US-08-961-527-181	Sequence 181, App
c	849	15	0.7	8798	4	US-09-182-117-4	Sequence 4, Appli
c	850	15	0.7	8868	3	US-08-815-809-7	Sequence 7, Appli
c	851	15	0.7	8898	4	US-08-961-527-69	Sequence 69, Appl
	852	15	0.7	9171	2	US-08-629-001A-2	Sequence 2, Appli
	853	15	0.7	9171	4	US-08-642-274D-2	Sequence 2, Appli
	854	15	0.7	9171	4	US-08-952-127-2	Sequence 2, Appli
	855	15	0.7	9171	4	US-08-952-014C-2	Sequence 2, Appli
	856	15	0.7	9244	4	US-08-961-527-68	Sequence 68, Appl
	857	15	0.7	9385	2	US-08-874-266-1	Sequence 1, Appli
	858	15	0.7	9385	4	US-09-360-416-1	Sequence 1, Appli
c	859	15	0.7	9461	4	US-09-221-017B-513	Sequence 513, App
c	860	15	0.7	9515	1	US-08-920-812-13	Sequence 13, Appl
c	861	15	0.7	9515	1	US-08-920-827-13	Sequence 13, Appl
c	862	15	0.7	9515	1	US-08-921-177-13	Sequence 13, Appl
c	863	15	0.7	9515	1	US-08-362-577C-13	Sequence 13, Appl
c	864	15	0.7	9515	2	US-08-920-828-13	Sequence 13, Appl
c	865	15	0.7	9713	4	US-08-961-527-43	Sequence 43, Appl
	866	15	0.7	9870	1	US-08-508-836A-9	Sequence 9, Appli
c	867	15	0.7	9897	4	US-08-961-527-10	Sequence 10, Appl
c	868	15	0.7	10299	2	US-08-477-451-1	Sequence 1, Appli
	869	15	0.7	10357	4	US-08-961-527-191	Sequence 191, App
	870	15	0.7	10395	1	US-08-245-809-3	Sequence 3, Appli
	871	15	0.7	10396	1	US-08-245-809-5	Sequence 5, Appli
c	872	15	0.7	10754	2	US-08-966-958-1	Sequence 1, Appli
c	873	15	0.7	10754	2	US-09-215-817-1	Sequence 1, Appli
c	874	15	0.7	10754	4	US-09-342-353-1	Sequence 1, Appli
	875	15	0.7	10798	1	US-08-107-748-2	Sequence 2, Appli
	876	15	0.7	10798	5	PCT-US92-01385-2	Sequence 2, Appli
c	877	15	0.7	10846	4	US-09-098-219B-5	Sequence 5, Appli
c	878	15	0.7	10900	4	US-09-098-219B-6	Sequence 6, Appli
	879	15	0.7	10965	1	US-08-107-748-4	Sequence 4, Appli
	880	15	0.7	10965	5	PCT-US92-01385-4	Sequence 4, Appli
	881	15	0.7	11091	4	US-09-134-001C-2243	Sequence 2243, Ap
	882	15	0.7	11461	3	US-08-669-161A-29	Sequence 29, Appl
	883	15	0.7	11770	4	US-08-961-527-172	Sequence 172, App
	884	15	0.7	12666	4	US-08-961-527-137	Sequence 137, App
c	885	15	0.7	13011	2	US-08-791-849A-14	Sequence 14, Appl
	886	15	0.7	13158	2	US-08-687-080-105	Sequence 105, App
	887	15	0.7	17656	4	US-09-433-579-3	Sequence 3, Appli
	888	15	0.7	18475	4	US-08-961-527-38	Sequence 38, Appl
c	889	15	0.7	18609	4	US-08-943-731-1	Sequence 1, Appli
c	890	15	0.7	19011	1	US-08-310-356-36	Sequence 36, Appl
c	891	15	0.7	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	892	15	0.7	19250	4	US-08-961-527-35	Sequence 35, Appl
c	893	15	0.7	19390	4	US-08-961-527-86	Sequence 86, Appl
c	894	15	0.7	19557	5	PCT-US92-06300-1	Sequence 1, Appli
c	895	15	0.7	19877	2	US-08-816-155B-8	Sequence 8, Appli
c	896	15	0.7	19877	3	US-09-079-587-8	Sequence 8, Appli
c	897	15	0.7	20137	4	US-09-262-773-206	Sequence 206, App
c	898	15	0.7	20138	4	US-09-262-773-9	Sequence 9, Appli
	899	15	0.7	21126	1	US-08-008-216-19	Sequence 19, Appl
c	900	15	0.7	21126	1	US-08-008-216-19	Sequence 19, Appl
	901	15	0.7	21126	1	US-08-459-569-19	Sequence 19, Appl
c	902	15	0.7	21126	1	US-08-459-569-19	Sequence 19, Appl
	903	15	0.7	21126	1	US-08-458-831-19	Sequence 19, Appl
c	904	15	0.7	21126	1	US-08-458-831-19	Sequence 19, Appl
c	905	15	0.7	23071	4	US-09-262-773-210	Sequence 210, App
c	906	15	0.7	28720	4	US-09-341-587-7	Sequence 7, Appli
c	907	15	0.7	31728	4	US-09-453-702B-64	Sequence 64, Appl
c	908	15	0.7	38844	4	US-09-734-675-3	Sequence 3, Appli
c	909	15	0.7	45546	4	US-09-146-053-6	Sequence 6, Appli
	910	15	0.7	45716	4	US-08-965-048-5	Sequence 5, Appli
	911	15	0.7	45989	4	US-08-965-048-6	Sequence 6, Appli
c	912	15	0.7	49136	4	US-09-422-869-1	Sequence 1, Appli
	913	15	0.7	50000	4	US-09-146-053-3	Sequence 3, Appli
c	914	15	0.7	50000	4	US-09-146-053-3	Sequence 3, Appli
c	915	15	0.7	50000	4	US-09-146-053-4	Sequence 4, Appli
	916	15	0.7	51259	3	US-08-781-891-209	Sequence 209, App
	917	15	0.7	51952	3	US-08-947-823-1	Sequence 1, Appli
	918	15	0.7	59065	4	US-09-813-817-3	Sequence 3, Appli
	919	15	0.7	59065	4	US-09-978-197-3	Sequence 3, Appli
c	920	15	0.7	70000	4	US-09-851-896-3	Sequence 3, Appli
	921	15	0.7	72604	4	US-09-268-992-7	Sequence 7, Appli
	922	15	0.7	72604	4	US-09-657-474-7	Sequence 7, Appli
	923	15	0.7	72928	3	US-09-009-913-1	Sequence 1, Appli
c	924	15	0.7	80246	4	US-09-078-294-4	Sequence 4, Appli
c	925	15	0.7	80595	4	US-09-078-294-3	Sequence 3, Appli
c	926	15	0.7	81001	4	US-09-750-580-1	Sequence 1, Appli
c	927	15	0.7	84495	4	US-09-797-906-3	Sequence 3, Appli
c	928	15	0.7	87350	3	US-08-781-891-79	Sequence 79, Appl
c	929	15	0.7	87543	4	US-09-791-211-3	Sequence 3, Appli
	930	15	0.7	99500	4	US-09-798-096-10	Sequence 10, Appl
c	931	15	0.7	112132	4	US-09-741-150-3	Sequence 3, Appli
c	932	15	0.7	152331	3	US-09-128-155-16	Sequence 16, Appl
c	933	15	0.7	162450	4	US-09-345-882-1	Sequence 1, Appli
	934	15	0.7	168575	4	US-09-426-290-1	Sequence 1, Appli
	935	15	0.7	169998	4	US-09-676-610B-24	Sequence 24, Appl
c	936	15	0.7	176373	3	US-09-128-155-17	Sequence 17, Appl
	937	15	0.7	246240	2	US-08-724-394A-20	Sequence 20, Appl
	938	15	0.7	246240	2	US-08-724-394A-21	Sequence 21, Appl
	939	15	0.7	246240	2	US-08-724-394A-22	Sequence 22, Appl
	940	15	0.7	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	941	15	0.7	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	942	14	0.6	14	1	US-08-298-073-29	Sequence 29, Appl
c	943	14	0.6	14	1	US-08-794-153-29	Sequence 29, Appl
c	944	14	0.6	14	4	US-09-115-566-29	Sequence 29, Appl
	945	14	0.6	18	3	US-08-181-664-16	Sequence 16, Appl
c	946	14	0.6	20	1	US-08-089-996-4	Sequence 4, Appli
c	947	14	0.6	20	2	US-08-478-178A-4	Sequence 4, Appli
c	948	14	0.6	20	2	US-08-488-177-4	Sequence 4, Appli
c	949	14	0.6	20	2	US-08-481-072A-4	Sequence 4, Appli
c	950	14	0.6	20	2	US-08-664-336-4	Sequence 4, Appli
c	951	14	0.6	20	2	US-08-481-066A-4	Sequence 4, Appli
c	952	14	0.6	20	3	US-08-578-615A-4	Sequence 4, Appli
c	953	14	0.6	20	4	US-09-225-749-4	Sequence 4, Appli
c	954	14	0.6	20	4	US-08-829-637A-4	Sequence 4, Appli
c	955	14	0.6	20	4	US-09-254-322-4	Sequence 4, Appli
c	956	14	0.6	20	5	PCT-US93-02213-4	Sequence 4, Appli
c	957	14	0.6	20	5	PCT-US94-07770-4	Sequence 4, Appli
	958	14	0.6	21	4	US-09-593-012-92	Sequence 92, Appl
	959	14	0.6	21	4	US-09-593-012-186	Sequence 186, App
c	960	14	0.6	22	3	US-09-255-893-2	Sequence 2, Appli
c	961	14	0.6	24	1	US-07-741-940-94	Sequence 94, Appl
c	962	14	0.6	24	1	US-08-289-548A-94	Sequence 94, Appl
c	963	14	0.6	24	1	US-08-452-654-94	Sequence 94, Appl
c	964	14	0.6	24	1	US-08-452-655B-94	Sequence 94, Appl
c	965	14	0.6	24	2	US-08-117-952-596	Sequence 596, App
c	966	14	0.6	24	3	US-08-450-582-94	Sequence 94, Appl
c	967	14	0.6	24	4	US-08-449-731-94	Sequence 94, Appl
c	968	14	0.6	27	4	US-08-584-040-5109	Sequence 5109, Ap
	969	14	0.6	28	2	US-08-705-196-5	Sequence 5, Appli
	970	14	0.6	30	3	US-08-448-619-1	Sequence 1, Appli
c	971	14	0.6	31	3	US-08-843-409-4	Sequence 4, Appli
	972	14	0.6	31	3	US-08-843-409-5	Sequence 5, Appli
c	973	14	0.6	34	1	US-08-476-634-9	Sequence 9, Appli
c	974	14	0.6	34	1	US-08-484-518-9	Sequence 9, Appli
c	975	14	0.6	34	1	US-08-943-834-9	Sequence 9, Appli
	976	14	0.6	38	4	US-09-482-607-2	Sequence 2, Appli

```
   977      14    0.6     40  1  US-08-443-957-15         Sequence 15, Appl
   978      14    0.6     42  4  US-08-471-112A-14        Sequence 14, Appl
   979      14    0.6     44  2  US-08-823-516-115        Sequence 115, App
   980      14    0.6     44  3  US-08-759-038-123        Sequence 123, App
   981      14    0.6     44  3  US-08-758-314-123        Sequence 123, App
   982      14    0.6     45  2  US-08-188-583-43         Sequence 43, Appl
 c 983      14    0.6     45  2  US-08-680-326-68         Sequence 68, Appl
 c 984      14    0.6     45  3  US-08-654-623-12         Sequence 12, Appl
 c 985      14    0.6     46  4  US-09-313-221A-56        Sequence 56, Appl
   986      14    0.6     47  2  US-08-188-583-48         Sequence 48, Appl
   987      14    0.6     47  2  US-08-332-562A-64        Sequence 64, Appl
   988      14    0.6     50  2  US-08-823-516-118        Sequence 118, App
   989      14    0.6     50  3  US-08-759-038-126        Sequence 126, App
   990      14    0.6     50  3  US-08-758-314-126        Sequence 126, App
   991      14    0.6     52  1  US-08-418-859-45         Sequence 45, Appl
   992      14    0.6     52  2  US-08-643-181-45         Sequence 45, Appl
 c 993      14    0.6     52  4  US-09-485-737B-100       Sequence 100, App
 c 994      14    0.6     54  4  US-09-449-218D-37        Sequence 37, Appl
   995      14    0.6     55  4  US-09-485-737B-99        Sequence 99, Appl
   996      14    0.6     56  2  US-08-190-199A-55        Sequence 55, Appl
   997      14    0.6     56  2  US-08-190-199A-57        Sequence 57, Appl
 c 998      14    0.6     60  3  US-08-654-623-14         Sequence 14, Appl
 c 999      14    0.6     67  4  US-09-485-737B-97        Sequence 97, Appl
  1000      14    0.6     70  1  US-07-974-251-2          Sequence 2, Appli

                                   ALIGNMENTS

RESULT 1
US-09-058-489-90
; Sequence 90, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 90
;   LENGTH: 2319
;   TYPE: DNA
;   ORGANISM: Human
US-09-058-489-90

  Query Match             0.9%;  Score 19;  DB 3;  Length 2319;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1830 ACATGTTATCAATTTTGAT 1848
              |||||||||||||||||||
Db       1598 ACATGTTATCAATTTTGAT 1616

RESULT 2
US-09-184-964-2
; Sequence 2, Application US/09184964
; Patent No. 6391574
;  GENERAL INFORMATION:
;    APPLICANT:  Rine, Jasper D.
;    APPLICANT:  Boyartchuk, Victor L.
;    APPLICANT:  Ashby, Matthew N.
;    TITLE OF INVENTION:  AFC1 AND RCE1: ISOPRENYLATED CAAX
;    TITLE OF INVENTION:  PROCESSING ENZYMES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  2200 Sand Hill road, suite 100
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/184,964
;      FILING DATE:  03-NOV-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/902,774
;      FILING DATE:  30-JUL-1997
;      APPLICATION NUMBER:  60/023,491
;      FILING DATE:  07-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Suyat, Reginald J.
;      REGISTRATION NUMBER:  28,172
;      REFERENCE/DOCKET NUMBER:  09272-006004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/322-5070
;      TELEFAX:  650/854-0875
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2850 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  Genomic DNA
US-09-184-964-2

  Query Match             0.9%;  Score 19;  DB 4;  Length 2850;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1674 TTTTCTTTGTCAAGAAAAA 1692
              |||||||||||||||||||
Db        919 TTTTCTTTGTCAAGAAAAA 937

RESULT 3
US-09-058-489-14
; Sequence 14, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
;   LENGTH: 3408
;   TYPE: DNA
;   ORGANISM: Human
US-09-058-489-14

  Query Match             0.9%;  Score 19;  DB 3;  Length 3408;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1829 AACATGTTATCAATTTTGA 1847
              |||||||||||||||||||
```

```
Db       2388 AACATGTTATCAATTTTGA 2406

RESULT 4
US-09-058-489-17
; Sequence 17, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
;   LENGTH: 4416
;   TYPE: DNA
;   ORGANISM: Human
US-09-058-489-17

  Query Match             0.9%;  Score 19;  DB 3;  Length 4416;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1830 ACATGTTATCAATTTTGAT 1848
              |||||||||||||||||||
Db       1598 ACATGTTATCAATTTTGAT 1616

RESULT 5
US-09-058-489-13
; Sequence 13, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 5322
;   TYPE: DNA
;   ORGANISM: Human
US-09-058-489-13

  Query Match             0.9%;  Score 19;  DB 3;  Length 5322;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1829 AACATGTTATCAATTTTGA 1847
              |||||||||||||||||||
Db       2388 AACATGTTATCAATTTTGA 2406

RESULT 6
US-08-651-155B-116
; Sequence 116, Application US/08651155B
; Patent No. 6365401
;  GENERAL INFORMATION:
;    APPLICANT:  Mahan Dr., Michael J.
;    APPLICANT:  Conner Mr., Christopher P.
;    APPLICANT:  Hiethoff Mr., Douglas M.
;    TITLE OF INVENTION:  METHOD AND PROBES FOR THE IDENTIFICATION
;    TITLE OF INVENTION:  OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST
;    TITLE OF INVENTION:  INFECTION
;    NUMBER OF SEQUENCES:  255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chrisman, Bynum & Johnson, P.C.
;      STREET:  1900 Fifteenth Street
;      CITY:  Boulder
;      STATE:  CO
;      COUNTRY:  USA
;      ZIP:  80302
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/651,155B
;      FILING DATE:  17-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Petersen Mr., Steven C.
;      REGISTRATION NUMBER:  36,238
;      REFERENCE/DOCKET NUMBER:  17060.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303/546-1300
;      TELEFAX:  303/449-5426
;      TELEX:  ABA1475
;  INFORMATION FOR SEQ ID NO:  116:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  339 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-651-155B-116

  Query Match             0.8%;  Score 18;  DB 4;  Length 339;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        110 CTCCAGCTTCATCATCAG 127
              ||||||||||||||||||
Db        184 CTCCAGCTTCATCATCAG 201

RESULT 7
US-09-222-575-116
; Sequence 116, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
;  APPLICANT: Yuqiu, Jiang
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
;  TITLE OF INVENTION:  and Methods for Their Use
;  FILE REFERENCE: 210121.470
;  CURRENT APPLICATION NUMBER: US/09/222,575
;  CURRENT FILING DATE:  1998-12-28
;  NUMBER OF SEQ ID NOS: 174
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 116
;   LENGTH: 421
;   TYPE: DNA
;   ORGANISM: Human
US-09-222-575-116

  Query Match             0.8%;  Score 18;  DB 4;  Length 421;
```

```
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1369 GGTCCAGAAATGAAGAAG 1386
              ||||||||||||||||||
Db         12 GGTCCAGAAATGAAGAAG 29


RESULT 8
US-09-318-443-3
; Sequence 3, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
;  APPLICANT: Hemmati-Brivanlou, Ali
;  APPLICANT: Weinstein, Daniel C.
;  TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: 600-1-211 N
;  CURRENT APPLICATION NUMBER: US/09/318,443
;  CURRENT FILING DATE: 1999-05-25
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 532
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-318-443-3

  Query Match             0.8%;  Score 18;  DB 4;  Length 532;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1330 GTTTTGGATGAAGCTGAT 1347
              ||||||||||||||||||
Db        346 GTTTTGGATGAAGCTGAT 363


RESULT 9
US-09-318-443-5
; Sequence 5, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
;  APPLICANT: Hemmati-Brivanlou, Ali
;  APPLICANT: Weinstein, Daniel C.
;  TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: 600-1-211 N
;  CURRENT APPLICATION NUMBER: US/09/318,443
;  CURRENT FILING DATE: 1999-05-25
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 1536
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-318-443-5

  Query Match             0.8%;  Score 18;  DB 4;  Length 1536;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1330 GTTTTGGATGAAGCTGAT 1347
              ||||||||||||||||||
Db        626 GTTTTGGATGAAGCTGAT 643


RESULT 10
US-09-318-443-7
; Sequence 7, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
;  APPLICANT: Hemmati-Brivanlou, Ali
;  APPLICANT: Weinstein, Daniel C.
;  TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: 600-1-211 N
;  CURRENT APPLICATION NUMBER: US/09/318,443
;  CURRENT FILING DATE: 1999-05-25
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 1682
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-318-443-7

  Query Match             0.8%;  Score 18;  DB 4;  Length 1682;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1330 GTTTTGGATGAAGCTGAT 1347
              ||||||||||||||||||
Db        767 GTTTTGGATGAAGCTGAT 784


RESULT 11
US-09-255-392-1/c
; Sequence 1, Application US/09255392
; Patent No. 6214850
;  GENERAL INFORMATION:
;    APPLICANT:  Evans, Ronald M.
;    APPLICANT:  Forman, Barry M.
;    TITLE OF INVENTION:  MODULATORS OF PEROXISOME PROLIFERATOR
;    TITLE OF INVENTION:  ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE USE THEREOF
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/255,392
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,493
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P41 9958
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-1995
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2005 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  352..1776
US-09-255-392-1

  Query Match             0.8%;  Score 18;  DB 4;  Length 1796;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        760 GATGAGGACTCCATCTTT 777
              ||||||||||||||||||
Db       1368 GATGAGGACTCCATCTTT 1351

RESULT 12
US-09-082-092-5
; Sequence 5, Application US/09082092
; Patent No. 6251628
;  GENERAL INFORMATION:
;    APPLICANT:  Nakao, Atsuhito
;    APPLICANT:  Moren, Anita
;    APPLICANT:  Heuchel, Rainer
;    APPLICANT:  Itoh, Susumu
;    APPLICANT:  Afrakhte, Mozghan
;    APPLICANT:  Souchelnytskyi, Serhiy
;    APPLICANT:  Brodin, Greger
;    APPLICANT:  Landstrom, Marene
;    APPLICANT:  Heldin, Nils-Erik
;    APPLICANT:  Heldin, Carl-Henrik
;    APPLICANT:  ten Dijke, Peter
;    TITLE OF INVENTION:  SMAD7 AND USES THEREOF
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/082,092
;      FILING DATE:  20-MAY-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/047,221
;      FILING DATE:  20-MAY-1997
;      APPLICATION NUMBER:  60/060,465
;      FILING DATE:  30-SEP-1997
;      APPLICATION NUMBER:  60/075,940
;      FILING DATE:  25-FEB-1998
;      APPLICATION NUMBER:  60/077,033
;      FILING DATE:  06-MAR-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Van Amsterdam, John R.
;      REGISTRATION NUMBER:  40,212
;      REFERENCE/DOCKET NUMBER:  L0461/7032
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1876 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  50...1327
;      OTHER INFORMATION:
US-09-082-092-5

  Query Match             0.8%;  Score 18;  DB 4;  Length 1876;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        631 TACAAAGGTTTAAATGAA 648
              ||||||||||||||||||
Db       1503 TACAAAGGTTTAAATGAA 1520

RESULT 13
US-09-291-922-19
; Sequence 19, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT: Hitz, Bill
;  APPLICANT: Kinney, Tony
;  APPLICANT: Tingey, Scott
;  TITLE OF INVENTION: Plant Sugar Transport Proteins
;  FILE REFERENCE: BB-1163
;  CURRENT APPLICATION NUMBER: US/09/291,922
;  CURRENT FILING DATE: 1999-04-14
;  EARLIER APPLICATION NUMBER: 60/083,044
;  EARLIER FILING DATE: April 24, 1998
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
;   LENGTH: 1914
;   TYPE: DNA
;   ORGANISM: Zea mays
US-09-291-922-19

  Query Match             0.8%;  Score 18;  DB 4;  Length 1914;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2121 TGGGTTTTCTTCTTCACG 2138
              ||||||||||||||||||
Db       1445 TGGGTTTTCTTCTTCACG 1462

RESULT 14
US-08-484-200-1/c
; Sequence 1, Application US/08484200
; Patent No. 5861274
;  GENERAL INFORMATION:
;    APPLICANT:  EVANS, RONALD M.
;    APPLICANT:  FORMAN, BARRY M.
;    APPLICANT:  KLIEWER, STEVEN A.
;    APPLICANT:  ONG, ESTELITA S.
;    TITLE OF INVENTION:  NOVEL MEMBERS OF THE STEROID/THYROID
;    TITLE OF INVENTION:  SUPERFAMILY AND USES THEREFOR
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
;      STREET:  444 SOUTH FLOWER STREET, SUITE 2000
;      CITY:  LOS ANGELES
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,200
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  REITER, STEPHEN E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P41 9971
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-1995
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  1:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2005 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  352..1776
US-08-484-200-1

  Query Match             0.8%;  Score 18;  DB 2;  Length 2005;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        760 GATGAGGACTCCATCTTT 777
              ||||||||||||||||||
Db       1368 GATGAGGACTCCATCTTT 1351


RESULT 15
US-08-477-493-1/c
; Sequence 1, Application US/08477493
; Patent No. 5939442
;  GENERAL INFORMATION:
;    APPLICANT:  Evans, Ronald M.
;    APPLICANT:  Forman, Barry M.
;    TITLE OF INVENTION:  MODULATORS OF PEROXISOME PROLIFERATOR
;    TITLE OF INVENTION:  ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE USE THEREOF
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,493
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P41 9958
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-1995
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2005 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  352..1776
US-08-477-493-1

  Query Match             0.8%;  Score 18;  DB 2;  Length 2005;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        760 GATGAGGACTCCATCTTT 777
              ||||||||||||||||||
Db       1368 GATGAGGACTCCATCTTT 1351
```

Search completed: June 10, 2003, 14:53:24
Job time : 155 secs

```
OM nucleic - nucleic search, using sw model

Run on:         June 10, 2003, 10:40:24 ; Search time 487 Seconds
                                          (without alignments)
                                          10043.820 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  2172
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched:       2185239 seqs, 1125999159 residues

Word size :     0

Total number of hits satisfying chosen parameters:     4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :       N_Geneseq_101002:*
             1:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
             2:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
             3:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
             4:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
             5:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
             6:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
             7:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
             8:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
             9:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
            10:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
            11:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
            12:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
            13:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
            14:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
            15:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
            16:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
            17:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
            18:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
            19:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
            20:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
            21:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
            22:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
            23:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
            24:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2172	100.0	2224	22	AAD06354	Human vasa cDNA.
	2	579	26.7	1006	21	AAA44726	Human secreted exp
	3	370	17.0	677	22	AAS40979	cDNA encoding nove
	4	255	11.7	591	22	AAS41534	cDNA encoding nove
	5	255	11.7	591	22	AAL01169	Human reproductive
	6	255	11.7	591	23	ABL96630	Human testicular a
	7	220	10.1	2194	22	AAL04793	Human reproductive
	8	220	10.1	2194	23	ABL97687	Human testicular a
	9	220	10.1	2195	22	AAL04794	Human reproductive
	10	220	10.1	2195	23	ABL97688	Human testicular a
	11	220	10.1	2197	22	AAL04795	Human reproductive
	12	220	10.1	2197	23	ABL97689	Human testicular a
	13	60	2.8	60	24	ABN42617	Human spliced tran
	14	23	1.1	2187	24	ABI99243	Mouse ischaemic co
	15	21	1.0	1842	23	AAS53256	Haemophilus influe
	16	21	1.0	1830121	17	AAT42063	Haemophilus influe
	17	20	0.9	20	22	AAD06359	RT-PCR primer #2,
	18	20	0.9	2320	21	AAC37735	Arabidopsis thalia
	19	20	0.9	9832	24	ABL32657	Human immune syste
	20	20	0.9	2944528	24	ABA03041	Listeria monocytog
c	21	19	0.9	28	22	AAD06357	PCR primer #2, to
c	22	19	0.9	138	22	ABA72946	Human foetal liver
c	23	19	0.9	138	22	AAK21377	Human brain expres
c	24	19	0.9	138	22	AAK47539	Human bone marrow
c	25	19	0.9	138	22	AAI53370	Probe #22056 used
c	26	19	0.9	138	24	ABS21638	Human genome-deriv
c	27	19	0.9	592	22	ABA60407	Human foetal liver
c	28	19	0.9	592	22	AAK08685	Human brain expres
c	29	19	0.9	592	22	AAK34573	Human bone marrow
c	30	19	0.9	592	22	AAI40291	Probe #8977 used t
c	31	19	0.9	592	24	ABS09206	Human genome-deriv
	32	19	0.9	760	22	AAH03354	Human cDNA clone (
	33	19	0.9	811	23	ABL16909	Drosophila melanog
	34	19	0.9	851	22	AAH05178	Human cDNA clone (
	35	19	0.9	866	22	AAH05374	Human cDNA clone (
	36	19	0.9	1122	22	AAH01654	Candida norvegica
	37	19	0.9	1233	21	AAA29206	Human topoisomeras
	38	19	0.9	1290	22	AAH00497	Pichia anomala nuc
	39	19	0.9	1291	22	AAH00498	Pichia anomala nuc
c	40	19	0.9	1472	21	AAC98050	Human colon cancer
c	41	19	0.9	1479	22	AAH90759	CFE 63 coding sequ
	42	19	0.9	1898	22	AAH14078	Human cDNA sequenc
	43	19	0.9	1987	21	AAC77791	Human cancer assoc
c	44	19	0.9	2123	22	AAH13995	Human cDNA sequenc
c	45	19	0.9	2227	22	AAK94365	Human full-length
c	46	19	0.9	2262	24	ABK34739	Human cDNA for nov
c	47	19	0.9	2293	22	AAH15834	Human cDNA sequenc
c	48	19	0.9	2306	22	AAH17861	Human cDNA sequenc
	49	19	0.9	2487	22	AAH15763	Human cDNA sequenc
c	50	19	0.9	2592	23	AAS68253	DNA encoding novel
	51	19	0.9	2722	24	AAD24414	Human RNA metaboli
	52	19	0.9	2850	19	AAV17602	Yeast RCE1 gene en
c	53	19	0.9	3039	23	AAS68807	DNA encoding novel
c	54	19	0.9	3039	23	AAS89721	DNA encoding novel
c	55	19	0.9	3041	22	AAH17625	Human cDNA sequenc
	56	19	0.9	3123	22	AAH14548	Human cDNA sequenc
	57	19	0.9	3408	19	AAV69632	Dead Box X (DBX) g
c	58	19	0.9	3550	19	AAV44295	Human secreted pro
c	59	19	0.9	3550	22	AAF98469	Human cDNA clone C
c	60	19	0.9	3764	23	ABL16908	Drosophila melanog
c	61	19	0.9	3862	21	AAD01233	Human semaphorin 6
	62	19	0.9	4416	19	AAV69633	Dead Box Y (DBY) g
	63	19	0.9	5321	21	AAA29207	Human DBX1 coding
	64	19	0.9	5322	19	AAV69631	Dead Box X (DBX) g
	65	19	0.9	5526	23	AAS76845	DNA encoding novel
	66	19	0.9	5778	22	AAK53086	Human polynucleoti
	67	19	0.9	5804	22	AAK52102	Human polynucleoti
c	68	19	0.9	6060	22	AAH18729	Human cDNA sequenc
c	69	19	0.9	10753	22	AAC85975	Domestic Pig endog
c	70	19	0.9	26781	23	ABL01850	Drosophila melanog
c	71	19	0.9	35498	23	ABL20316	Drosophila melanog
c	72	19	0.9	50000	24	ABA98944	Human asthma-assoc
	73	19	0.9	534720	19	AAV30458	Rhizobium species
	74	19	0.9	536165	19	AAV30459	Rhizobium species
c	75	18	0.8	98	21	AAC28944	Human secreted pro
c	76	18	0.8	98	21	AAZ42502	Human 5' EST isola
	77	18	0.8	108	23	ABV53358	Human prostate exp
c	78	18	0.8	174	22	AAF88278	H. tuberculata HtH
	79	18	0.8	333	21	AAC28563	Human secreted pro
c	80	18	0.8	358	22	AAL24860	Human breast cance
c	81	18	0.8	379	21	AAC94751	Cat flea hindgut a
c	82	18	0.8	381	21	AAC94230	Cat flea head and

	83	18	0.8	421	22	AAS47061	Human breast cance
	84	18	0.8	421	22	AAF17631	Human breast cance
	85	18	0.8	421	24	ABK95096	Human breast tumou
	86	18	0.8	423	22	ABA58227	Human foetal liver
	87	18	0.8	423	22	AAK06314	Human brain expres
	88	18	0.8	423	22	AAK31981	Human bone marrow
	89	18	0.8	423	22	AAI37842	Probe #6528 used t
	90	18	0.8	423	24	ABS06745	Human genome-deriv
	91	18	0.8	452	22	AAH01646	Aspergillus fumiga
	92	18	0.8	517	24	ABK94752	Neurodegenerative
	93	18	0.8	532	22	AAD20194	Human translation
	94	18	0.8	532	22	AAF86070	Human translation
	95	18	0.8	535	21	AAC36920	Arabidopsis thalia
c	96	18	0.8	587	21	AAA74304	Loblolly pine SSR
c	97	18	0.8	599	24	ABL63546	Breast cancer rela
	98	18	0.8	733	20	AAZ15038	Human gene express
	99	18	0.8	777	24	ABQ55513	Human ovarian anti
	100	18	0.8	800	22	AAK63791	Human immune/haema
c	101	18	0.8	924	23	AAS73663	DNA encoding novel
	102	18	0.8	929	22	ABA77019	Proliferative glom
c	103	18	0.8	957	22	AAH31733	Human olfactory re
	104	18	0.8	1018	22	AAH00689	Cunninghamella ber
c	105	18	0.8	1150	21	AAC42061	Arabidopsis thalia
c	106	18	0.8	1205	24	ABK65178	Arabidopsis cDNA e
	107	18	0.8	1277	22	AAH00470	Candida dubliniens
	108	18	0.8	1289	23	ABL06963	Drosophila melanog
	109	18	0.8	1338	23	ABL29731	Drosophila melanog
	110	18	0.8	1536	22	AAD20195	Human translation
	111	18	0.8	1536	22	AAF86071	Human translation
	112	18	0.8	1558	23	ABL23041	Drosophila melanog
	113	18	0.8	1576	22	AAS44681	Human full-length
	114	18	0.8	1682	22	AAD20196	Human translation
	115	18	0.8	1682	22	AAF86072	Human translation
	116	18	0.8	1682	24	ABK83618	Human cDNA differe
	117	18	0.8	1876	20	AAX15144	cDNA encoding huma
	118	18	0.8	1878	23	ABL15051	Drosophila melanog
	119	18	0.8	1914	24	ABK51971	Corn cDNA clone ce
c	120	18	0.8	2005	17	AAX25481	Human peroxisome p
c	121	18	0.8	2005	17	AAT16205	Peroxisome prolife
c	122	18	0.8	2005	20	AAX29814	Mouse peroxisome p
	123	18	0.8	2086	24	AAD38868	Mouse isocitrate d
	124	18	0.8	2091	24	ABA95077	Mouse IDPc encodin
	125	18	0.8	2111	23	AAS87257	DNA encoding novel
	126	18	0.8	2480	21	AAA64202	cDNA sequence enco
c	127	18	0.8	2541	23	AAS85610	DNA encoding novel
	128	18	0.8	2577	23	AAS83889	DNA encoding novel
c	129	18	0.8	2715	20	AAV80665	Human DNAX toll-li
c	130	18	0.8	2715	24	AAD26282	Human DNAX Toll li
c	131	18	0.8	2781	22	AAI59581	Human polynucleoti
c	132	18	0.8	2794	22	AAI59580	Human polynucleoti
c	133	18	0.8	3001	21	AAH51739	Chromosome 13q31-q
c	134	18	0.8	3060	22	AAF77098	Arabidopsis gene #
	135	18	0.8	3103	21	AAZ50708	Nucleotide sequenc
	136	18	0.8	3105	18	AAT94468	Human Fchd540 gene
	137	18	0.8	3111	22	AAF26660	Human Smad7 nucleo
	138	18	0.8	3111	24	ABK64389	Human benign prost
	139	18	0.8	3111	24	ABL66567	Lung cancer relate
c	140	18	0.8	3321	21	AAC81290	Mouse apoptosis-re
	141	18	0.8	3459	23	ABL06962	Drosophila melanog
c	142	18	0.8	3513	22	ABA82683	TRIO like gene SEQ
	143	18	0.8	4935	23	ABL15050	Drosophila melanog
	144	18	0.8	5110	23	ABL12038	Drosophila melanog
c	145	18	0.8	6416	23	ABL23040	Drosophila melanog
c	146	18	0.8	6497	23	ABL29730	Drosophila melanog
c	147	18	0.8	6763	21	AAC81309	Mouse FLASH protei
	148	18	0.8	7232	23	ABL24792	Drosophila melanog
	149	18	0.8	8530	22	AAS31089	Human diagnostic a
	150	18	0.8	9192	12	AAQ11943	Nucleotide sequenc
c	151	18	0.8	9551	22	AAS39928	Genomic sequence #
c	152	18	0.8	9551	22	AAK90372	Human digestive sy
	153	18	0.8	10786	23	ABL10396	Drosophila melanog
c	154	18	0.8	10951	23	ABL24488	Drosophila melanog
c	155	18	0.8	11120	21	AAD01336	Drosophila melanog

	156	18	0.8	13467	22	AAL05944	Human reproductive
	157	18	0.8	13467	22	AAS27670	DNA encoding novel
	158	18	0.8	13467	22	AAS27838	DNA encoding novel
	159	18	0.8	13467	22	AAK85539	Human immune/haema
	160	18	0.8	13467	23	ABL98508	Human testicular a
c	161	18	0.8	15558	22	AAS39926	Genomic sequence #
c	162	18	0.8	15558	22	AAK90370	Human digestive sy
c	163	18	0.8	18385	22	AAS28584	Genomic sequence #
c	164	18	0.8	18385	22	AAK79804	Human immune/haema
	165	18	0.8	19399	23	ABL06008	Drosophila melanog
	166	18	0.8	20020	22	AAS29216	Genomic sequence #
c	167	18	0.8	20020	22	AAL05535	Human reproductive
	168	18	0.8	349980	22	AAH41225	Pyrococcus abyssi
	169	18	0.8	349980	22	AAH41226	Pyrococcus abyssi
c	170	18	0.8	1664976	19	AAV21209	Methanococcus jann
	171	17	0.8	24	24	ABQ00895	Oligonucleotide ad
	172	17	0.8	24	24	ABQ05732	Oligonucleotide ad
c	173	17	0.8	24	24	ABQ05773	Oligonucleotide ad
c	174	17	0.8	37	23	ABK05404	Human NOGO G-Cleav
c	175	17	0.8	135	22	ABA74956	Human foetal liver
c	176	17	0.8	135	22	AAK49601	Human bone marrow
c	177	17	0.8	135	22	AAI26709	Probe #16642 for g
c	178	17	0.8	135	22	AAI55483	Probe #24169 used
c	179	17	0.8	135	24	ABS23086	Human genome-deriv
	180	17	0.8	155	22	ABA71193	Human foetal liver
	181	17	0.8	155	22	AAK19491	Human brain expres
	182	17	0.8	155	22	AAK45482	Human bone marrow
	183	17	0.8	155	22	AAI51427	Probe #20113 used
	184	17	0.8	155	24	ABS19750	Human genome-deriv
c	185	17	0.8	168	22	ABA49389	Human breast cell
c	186	17	0.8	168	22	ABA67299	Human foetal liver
c	187	17	0.8	168	22	ABA34392	Probe #12858 for g
c	188	17	0.8	168	22	AAK15736	Human brain expres
c	189	17	0.8	168	22	AAK41473	Human bone marrow
c	190	17	0.8	168	22	AAI22220	Probe #12153 for g
c	191	17	0.8	168	22	AAI47516	Probe #16202 used
c	192	17	0.8	168	22	AAI07919	Probe #7910 used t
c	193	17	0.8	168	24	ABS15478	Human genome-deriv
	194	17	0.8	266	23	AAS76380	DNA encoding novel
c	195	17	0.8	278	24	ABL74960	Corn tassel-derive
c	196	17	0.8	300	21	AAZ50480	Human dickkopf-1 h
c	197	17	0.8	305	22	AAK63278	Human immune/haema
	198	17	0.8	308	20	AAV87930	EST clone FG567.
c	199	17	0.8	318	22	AAH66401	C glutamicum codin
c	200	17	0.8	320	22	AAS39536	cDNA encoding nove
c	201	17	0.8	320	22	AAK88435	Human digestive sy
c	202	17	0.8	359	23	ABV07739	Human prostate exp
c	203	17	0.8	366	22	AAI90950	Human polynucleoti
	204	17	0.8	370	24	ABQ85491	Arabidopsis thalia
c	205	17	0.8	387	22	AAI88380	Human polynucleoti
c	206	17	0.8	391	22	AAK72558	Human immune/haema
c	207	17	0.8	393	22	AAK58359	Human immune/haema
	208	17	0.8	416	14	AAQ60149	Human brain Expres
	209	17	0.8	420	16	AAT23577	Human gene signatu
c	210	17	0.8	420	23	ABV37667	Human prostate exp
c	211	17	0.8	421	24	ABL83321	Human ovarian canc
c	212	17	0.8	422	22	ABA44235	Human breast cell
c	213	17	0.8	422	22	ABA54683	Human foetal liver
c	214	17	0.8	422	22	ABA24465	Probe #2931 for ge
c	215	17	0.8	422	22	AAK02970	Human brain expres
c	216	17	0.8	422	22	AAK28415	Human bone marrow
c	217	17	0.8	422	22	AAI12982	Probe #2915 for ge
c	218	17	0.8	422	22	AAI34340	Probe #3026 used t
c	219	17	0.8	422	22	AAI02900	Probe #2891 used t
c	220	17	0.8	422	24	ABS02923	Human genome-deriv
c	221	17	0.8	448	22	AAI89681	Human polynucleoti
c	222	17	0.8	461	23	ABV60656	Human prostate exp
c	223	17	0.8	464	24	ABL87688	Human ovarian canc
	224	17	0.8	471	22	ABA43572	Human breast cell
	225	17	0.8	471	22	ABA54024	Human foetal liver
	226	17	0.8	471	22	ABA23779	Probe #2245 for ge
	227	17	0.8	471	22	AAK02297	Human brain expres
	228	17	0.8	471	22	AAK27741	Human bone marrow

229	17	0.8	471	22	AAI12321	Probe #2254 for ge
230	17	0.8	471	22	AAI33678	Probe #2364 used t
231	17	0.8	471	22	AAI02237	Probe #2228 used t
232	17	0.8	471	24	ABS02194	Human genome-deriv
c 233	17	0.8	472	22	ABA52030	Human foetal liver
c 234	17	0.8	472	22	ABA56987	Human foetal liver
c 235	17	0.8	472	22	ABA21844	Probe #310 for gen
c 236	17	0.8	472	22	AAK00314	Human brain expres
c 237	17	0.8	472	22	AAK25756	Human bone marrow
c 238	17	0.8	472	22	AAI10385	Probe #318 for gen
c 239	17	0.8	472	22	AAI15216	Probe #5149 for ge
c 240	17	0.8	472	22	AAI31637	Probe #323 used to
c 241	17	0.8	472	22	AAI00320	Probe #311 used to
c 242	17	0.8	472	24	ABS00336	Human genome-deriv
243	17	0.8	476	21	AAC35898	Arabidopsis thalia
244	17	0.8	476	22	ABA58695	Human foetal liver
245	17	0.8	476	22	AAK06830	Human brain expres
246	17	0.8	476	22	AAK32548	Human bone marrow
247	17	0.8	476	22	AAI38374	Probe #7060 used t
248	17	0.8	476	24	ABS07338	Human genome-deriv
249	17	0.8	483	18	AAV24801	H. pylori ORF 14cp
c 250	17	0.8	492	24	ABL36418	Human colon tumour
251	17	0.8	502	21	AAC44406	Arabidopsis thalia
c 252	17	0.8	512	22	ABA62431	Human foetal liver
c 253	17	0.8	512	22	AAK36647	Human bone marrow
c 254	17	0.8	512	22	AAI17500	Probe #7433 for ge
c 255	17	0.8	512	22	AAI42410	Probe #11096 used
c 256	17	0.8	512	24	ABS10645	Human genome-deriv
257	17	0.8	517	24	ABN61631	Human cancer relat
c 258	17	0.8	532	24	ABN63072	Human cancer relat
c 259	17	0.8	538	23	ABV08756	Human prostate exp
c 260	17	0.8	541	24	ABQ40246	Oligonucleotide fo
261	17	0.8	541	24	ABQ40247	Oligonucleotide fo
c 262	17	0.8	554	21	AAC40229	Arabidopsis thalia
263	17	0.8	555	18	AAV24914	H. pylori secreted
264	17	0.8	574	22	AAH10046	Human cDNA clone (
265	17	0.8	574	22	AAF93944	Primer specific fo
266	17	0.8	574	24	ABL36582	Human colon tumour
267	17	0.8	605	21	AAF14197	Aspergillus oryzae
268	17	0.8	616	23	ABV53546	Human prostate exp
c 269	17	0.8	624	21	AAA08843	Human DKR-2 gene s
c 270	17	0.8	626	23	ABV38653	Human prostate exp
271	17	0.8	629	23	AAS71047	DNA encoding novel
c 272	17	0.8	632	21	AAF08951	Fusarium venenatum
c 273	17	0.8	663	21	AAZ54144	Neisseria gonorrhe
274	17	0.8	697	20	AAX30352	DNA encoding a hum
c 275	17	0.8	702	19	AAV07909	Human cysteine-ric
c 276	17	0.8	702	21	AAA75134	cDNA encoding a pa
c 277	17	0.8	706	21	AAC36936	Arabidopsis thalia
278	17	0.8	719	21	AAC45206	Arabidopsis thalia
c 279	17	0.8	720	22	ABA07921	Human ovarian and
c 280	17	0.8	720	22	AAL03738	Human reproductive
c 281	17	0.8	720	22	AAK84566	Human immune/haema
282	17	0.8	722	21	AAA78467	Plant SDF polynucl
283	17	0.8	744	23	AAS55726	Streptococcus pneu
284	17	0.8	744	23	AAS55887	Streptococcus pneu
c 285	17	0.8	765	24	ABN66076	Streptococcus poly
c 286	17	0.8	765	24	ABN70280	Streptococcus poly
c 287	17	0.8	768	20	AAX56830	Human phdkk-2 cDNA
288	17	0.8	769	22	AAI95526	Human neuroblastom
c 289	17	0.8	778	18	AAX30821	Streptococcus pneu
c 290	17	0.8	780	21	AAA08842	Human DKR-2 DNA.
c 291	17	0.8	789	21	AAA75139	Open reading frame
292	17	0.8	795	21	AAC34416	Arabidopsis thalia
293	17	0.8	801	21	AAF07750	Fusarium venenatum
c 294	17	0.8	820	22	AAH07639	Human cDNA clone (
c 295	17	0.8	828	22	AAS39954	Genomic sequence #
c 296	17	0.8	828	22	AAS39955	Genomic sequence #
c 297	17	0.8	828	22	AAS39956	Genomic sequence #
c 298	17	0.8	828	22	AAS39957	Genomic sequence #
c 299	17	0.8	828	22	AAK90398	Human digestive sy
c 300	17	0.8	828	22	AAK90399	Human digestive sy
c 301	17	0.8	828	22	AAK90400	Human digestive sy
c 302	17	0.8	828	22	AAK90401	Human digestive sy
303	17	0.8	888	23	AAS93663	DNA encoding novel
304	17	0.8	909	22	AAH62810	Shrimp white spot
c 305	17	0.8	915	22	AAD16542	Human novel protei
c 306	17	0.8	960	23	AAS75149	DNA encoding novel
307	17	0.8	987	24	ABK74561	Bacillus lichenifo
c 308	17	0.8	1038	21	AAC46376	Arabidopsis thalia
c 309	17	0.8	1041	21	AAC49125	Arabidopsis thalia
310	17	0.8	1053	22	AAF32438	Human ribonucleoti
311	17	0.8	1053	22	AAF32447	Human ribonucleoti
312	17	0.8	1071	23	AAS91817	DNA encoding novel
313	17	0.8	1081	22	AAF32440	Human ribonucleoti
c 314	17	0.8	1089	21	AAA30581	Human G protein-co
c 315	17	0.8	1089	21	AAA30710	DNA encoding human
c 316	17	0.8	1089	24	AAD26832	Human G-protein co
c 317	17	0.8	1125	21	AAC43034	Arabidopsis thalia
318	17	0.8	1134	7	AAN60428	Sequence encoding
319	17	0.8	1182	22	AAH65738	C glutamicum codin
c 320	17	0.8	1196	14	AAQ34594	Clone for TGF-beta
321	17	0.8	1206	22	AAH66931	C glutamicum codin
322	17	0.8	1237	8	AAN70314	Sequence encoding
323	17	0.8	1239	21	AAC43222	Arabidopsis thalia
324	17	0.8	1266	22	AAF71669	Corynebacterium gl
325	17	0.8	1305	21	AAC46486	Arabidopsis thalia
326	17	0.8	1338	9	AAN80040	Sequence encoding
c 327	17	0.8	1362	23	AAS66186	DNA encoding novel
c 328	17	0.8	1375	23	ABV24711	Human prostate exp
329	17	0.8	1377	21	AAC45674	Arabidopsis thalia
330	17	0.8	1381	21	AAC36272	Arabidopsis thalia
c 331	17	0.8	1389	15	AAQ70737	TATA-binding prote
332	17	0.8	1401	15	AAQ72189	Rat PACAP receptor
c 333	17	0.8	1446	22	AAD15870	Muskmelon 9-HPL cD
334	17	0.8	1453	18	AAT61724	Streptococcus pneu
335	17	0.8	1464	24	ABQ67922	Listeria monocytog
336	17	0.8	1464	24	ABQ69939	Listeria monocytog
337	17	0.8	1485	15	AAQ72191	Rat PACAP receptor
c 338	17	0.8	1517	20	AAZ20021	Soybean brittle-1
339	17	0.8	1525	21	AAC49500	Arabidopsis thalia
340	17	0.8	1527	23	AAS90226	DNA encoding novel
341	17	0.8	1528	21	AAC33857	Arabidopsis thalia
c 342	17	0.8	1539	21	AAC46496	Arabidopsis thalia
343	17	0.8	1548	21	AAC40278	Arabidopsis thalia
c 344	17	0.8	1560	21	AAC51381	Arabidopsis thalia
c 345	17	0.8	1575	21	AAC39261	Arabidopsis thalia
c 346	17	0.8	1578	17	AAT42223	Human TATA-binding
c 347	17	0.8	1578	18	AAT79597	TATA-binding prote
c 348	17	0.8	1578	23	ABL04223	Drosophila melanog
349	17	0.8	1601	22	AAH14924	Human cDNA sequenc
c 350	17	0.8	1602	21	AAC51388	Arabidopsis thalia
c 351	17	0.8	1617	19	AAV21658	AAV4 VP3 coat prot
352	17	0.8	1618	21	AAA94725	Soybean auxin tran
353	17	0.8	1640	23	AAS90225	DNA encoding novel
c 354	17	0.8	1648	22	AAH13852	Human cDNA sequenc
355	17	0.8	1666	21	AAC48754	Arabidopsis thalia
356	17	0.8	1670	21	AAC33813	Arabidopsis thalia
357	17	0.8	1674	22	AAL06174	Human reproductive
358	17	0.8	1674	22	AAL06177	Human reproductive
359	17	0.8	1674	23	ABL98739	Human testicular a
360	17	0.8	1674	23	ABL98742	Human testicular a
c 361	17	0.8	1700	23	AAS85309	DNA encoding novel
c 362	17	0.8	1711	21	AAC47230	Arabidopsis thalia
c 363	17	0.8	1715	22	AAD15894	Muskmelon 9-HPL cD
c 364	17	0.8	1757	21	AAC41794	Arabidopsis thalia
c 365	17	0.8	1800	19	AAV21657	AAV4 VP2 coat prot
c 366	17	0.8	1802	23	ABL20783	Drosophila melanog
c 367	17	0.8	1813	21	AAA64012	DNA encoding a 9-h
c 368	17	0.8	1813	21	AAZ49782	Cucumber hydropero
369	17	0.8	1846	23	AAS76381	DNA encoding novel
370	17	0.8	1869	15	AAQ72205	Rat PACAP receptor
371	17	0.8	1879	22	AAS05386	Rat caspase recrui
372	17	0.8	1894	22	ABA21440	Human nervous syst
373	17	0.8	1896	22	ABA21439	Human nervous syst
c 374	17	0.8	1906	19	AAV39012	CD30 ligand gene u

c	375	17	0.8	1906	24	ABK83816	Human cDNA differe
	376	17	0.8	1924	19	AAV27403	Streptococcus pneu
	377	17	0.8	1924	24	ABQ84871	S. pneumoniae SP09
c	378	17	0.8	1926	21	AAC36370	Arabidopsis thalia
c	379	17	0.8	1929	18	AAT72642	Human testis-speci
c	380	17	0.8	1929	20	AAX09015	Testis specific is
c	381	17	0.8	1930	21	AAC45681	Arabidopsis thalia
	382	17	0.8	2028	22	AAF57106	S. tuberosum CBP4
c	383	17	0.8	2060	23	ABL04259	Drosophila melanog
c	384	17	0.8	2120	21	AAZ56978	Arabidopsis E3-2 s
c	385	17	0.8	2127	23	AAS88963	DNA encoding novel
	386	17	0.8	2179	19	AAT98571	DNA encoding a S.
c	387	17	0.8	2208	19	AAV21650	AAV4 VP1 capsid pr
	388	17	0.8	2219	24	AAD33878	Human m32404 cDNA.
c	389	17	0.8	2234	23	ABL09662	Drosophila melanog
	390	17	0.8	2235	21	AAA05407	Streptococcus pneu
c	391	17	0.8	2256	24	ABN67495	Streptococcus poly
c	392	17	0.8	2256	24	ABN70389	Streptococcus poly
c	393	17	0.8	2291	22	AAI59445	Human polynucleoti
	394	17	0.8	2294	21	AAC63433	Human secreted pro
	395	17	0.8	2294	24	ABL89851	Human polynucleoti
c	396	17	0.8	2321	22	AAI58680	Human polynucleoti
c	397	17	0.8	2325	19	AAZ96293	S. pneumoniae deri
c	398	17	0.8	2326	22	AAH18611	Human cDNA sequenc
	399	17	0.8	2344	21	AAC38986	Arabidopsis thalia
	400	17	0.8	2348	21	AAC37726	Arabidopsis thalia
c	401	17	0.8	2366	22	AAI60510	Human polynucleoti
	402	17	0.8	2386	19	AAV35645	Cladosporium oxysp
c	403	17	0.8	2510	22	AAH75515	Human huntingtin i
	404	17	0.8	2548	15	AAQ72206	Rat PACAP receptor
	405	17	0.8	2559	23	AAS78444	DNA encoding novel
	406	17	0.8	2559	23	AAS87347	DNA encoding novel
c	407	17	0.8	2559	23	AAS89198	DNA encoding novel
	408	17	0.8	2560	23	AAS90227	DNA encoding novel
	409	17	0.8	2564	24	ABL58957	Human tumour marke
	410	17	0.8	2596	21	AAA12411	cDNA encoding a hu
c	411	17	0.8	2608	16	AAQ95541	Partial ALK gene.
	412	17	0.8	2649	15	AAQ55724	Alpha-DNA polymera
c	413	17	0.8	2715	24	ABN79809	Fungal ZBC gene se
c	414	17	0.8	2756	22	AAH42607	Nucleotide sequenc
c	415	17	0.8	2846	19	AAV20806	Homo sapiens vascu
c	416	17	0.8	2846	22	AAA91006	Human VEGF-D codin
	417	17	0.8	2894	21	AAC63431	Human secreted pro
	418	17	0.8	2927	23	ABL13202	Drosophila melanog
c	419	17	0.8	2932	24	AAD26831	Human G-protein co
c	420	17	0.8	2932	24	AAD26878	Human G-protein co
	421	17	0.8	2986	21	AAA23448	cDNA encoding huma
	422	17	0.8	3112	15	AAQ53948	Thermophilic DNA p
c	423	17	0.8	3115	23	ABL17416	Drosophila melanog
c	424	17	0.8	3117	22	AAH24264	Human phosphatase
	425	17	0.8	3124	22	AAF44826	RPP5-like protein
c	426	17	0.8	3170	21	AAA37042	Human PRO1316 (UNQ
c	427	17	0.8	3170	22	AAS46049	Human DNA encoding
c	428	17	0.8	3170	22	AAF54250	DNA encoding prote
	429	17	0.8	3393	20	AAX24681	Human synaptonemal
	430	17	0.8	3393	22	AAH76512	Human SCP-1 coding
c	431	17	0.8	3422	22	AAL03358	Human reproductive
c	432	17	0.8	3422	22	AAL03360	Human reproductive
c	433	17	0.8	3422	22	AAS28909	Human immunoglobul
	434	17	0.8	3422	22	AAS28911	Human immunoglobul
c	435	17	0.8	3551	22	AAH18414	Human cDNA sequenc
c	436	17	0.8	3587	23	ABL04840	Drosophila melanog
c	437	17	0.8	3696	21	AAA75138	cDNA encoding a hu
c	438	17	0.8	3718	24	ABK63564	Rat sequence diffe
	439	17	0.8	3748	23	ABL06888	Drosophila melanog
	440	17	0.8	3752	19	AAV52341	Streptococcus pneu
c	441	17	0.8	4066	19	AAV34331	Human glycogenin-2
c	442	17	0.8	4102	23	ABL28625	Drosophila melanog
	443	17	0.8	4285	22	AAI58272	Human polynucleoti
	444	17	0.8	4287	22	AAI60058	Human polynucleoti
c	445	17	0.8	4359	22	AAL35981	Human musculoskele
c	446	17	0.8	4360	22	AAL35982	Human musculoskele
	447	17	0.8	4367	23	AAS86538	DNA encoding novel

c	448	17	0.8	4494	22	AAD18129	Rat TGF-beta Resis
	449	17	0.8	4512	22	AAH16662	Human cDNA sequenc
	450	17	0.8	4746	23	ABL23508	Drosophila melanog
c	451	17	0.8	4767	19	AAV21648	AAV4 genome. Aden
	452	17	0.8	4948	22	AAK84811	Human immune/haema
	453	17	0.8	4955	22	AAF32439	Human ribonucleoti
c	454	17	0.8	4998	22	AAS01079	Mouse vitronectin
c	455	17	0.8	4998	24	AAD22527	Mouse vitronectin
	456	17	0.8	5021	22	AAK81320	Human immune/haema
c	457	17	0.8	5220	23	ABL28025	Drosophila melanog
	458	17	0.8	5489	24	ABL34120	Human immune syste
c	459	17	0.8	5850	20	AAX83476	Group B streptococ
	460	17	0.8	5881	22	AAL07230	Human reproductive
	461	17	0.8	5881	23	ABL98778	Human testicular a
	462	17	0.8	5930	24	ABL33394	Human immune syste
	463	17	0.8	6051	22	AAS46604	Tumour suppressor
	464	17	0.8	6051	24	ABL33851	Human immune syste
c	465	17	0.8	6135	23	ABV23953	Human prostate exp
c	466	17	0.8	6135	23	ABV24116	Human prostate exp
c	467	17	0.8	6135	23	ABV24962	Human prostate exp
c	468	17	0.8	6135	23	ABV25623	Human prostate exp
c	469	17	0.8	6135	23	ABV29836	Human prostate exp
c	470	17	0.8	6135	23	ABV29995	Human prostate exp
c	471	17	0.8	6226	22	AAF76867	Human ALK gene. H
	472	17	0.8	6582	18	AAT99451	S. carnosus nitrat
	473	17	0.8	6654	23	ABL20782	Drosophila melanog
	474	17	0.8	6853	23	ABL04258	Drosophila melanog
	475	17	0.8	7215	21	AAD00831	Hybrid adeno-assoc
	476	17	0.8	7228	23	ABL04222	Drosophila melanog
c	477	17	0.8	8151	21	AAD00832	Chimeric adeno-ass
c	478	17	0.8	8160	20	AAV99559	Arabidopsis lysine
c	479	17	0.8	8197	24	ABL70541	Chemically treated
c	480	17	0.8	8197	24	ABL34514	Human metastasis a
c	481	17	0.8	8440	23	ABL28024	Drosophila melanog
c	482	17	0.8	8512	19	AAV52262	Streptococcus pneu
	483	17	0.8	9453	21	AAF21116	Human low adenosin
	484	17	0.8	9453	21	AAA34994	Human adenosine re
	485	17	0.8	9453	24	ABK84194	Human cDNA differe
	486	17	0.8	9453	24	ABL65014	Lung cancer relate
c	487	17	0.8	10198	22	AAL06055	Human reproductive
c	488	17	0.8	10198	22	AAL06056	Human reproductive
c	489	17	0.8	10198	23	ABL98620	Human testicular a
c	490	17	0.8	10198	23	ABL98621	Human testicular a
c	491	17	0.8	10336	22	ABA17315	Human nervous syst
c	492	17	0.8	10514	23	ABL28624	Drosophila melanog
	493	17	0.8	10872	24	ABL70335	Chemically treated
	494	17	0.8	10872	24	AAS61288	Human gene regulat
	495	17	0.8	10872	24	ABK31388	Signal transductio
c	496	17	0.8	10910	23	ABL03131	Drosophila melanog
c	497	17	0.8	11389	19	AAV34334	Human glycogenin-2
	498	17	0.8	11735	22	AAS45329	Chemically pretrea
	499	17	0.8	11735	24	AAS61142	Human gene regulat
	500	17	0.8	11735	24	ABK28168	DNA transcription
	501	17	0.8	12025	24	ABL33299	Human immune syste
	502	17	0.8	12831	22	AAK70373	Human immune/haema
	503	17	0.8	13165	22	ABA07140	Human pancreatic c
	504	17	0.8	13165	22	AAK89348	Human digestive sy
c	505	17	0.8	14119	23	ABL17974	Drosophila melanog
	506	17	0.8	14800	24	ABL66291	Lung cancer relate
	507	17	0.8	14835	24	AAS94858	Human DNA sequence
	508	17	0.8	15240	21	AAF21117	Human low adenosin
	509	17	0.8	15240	21	AAA34995	Human adenosine re
	510	17	0.8	15387	24	ABL32184	Human immune syste
c	511	17	0.8	16838	24	AAS16687	DNA encoding super
	512	17	0.8	17512	23	ABL09034	Drosophila melanog
c	513	17	0.8	17594	24	ABL34026	Human immune syste
c	514	17	0.8	18234	22	AAH62719	Shrimp white spot
c	515	17	0.8	18571	23	ABL09756	Drosophila melanog
c	516	17	0.8	18574	23	ABL17988	Drosophila melanog
c	517	17	0.8	18574	23	ABL18034	Drosophila melanog
	518	17	0.8	18786	23	AAS59579	Propionibacterium
c	519	17	0.8	20813	22	ABA17314	Human nervous syst
	520	17	0.8	28344	23	ABL21640	Drosophila melanog

	521	17	0.8 31129	24	AAD36229	Human transporter
c	522	17	0.8 32154	22	AAL37068	Human musculoskele
c	523	17	0.8 32768	20	AAX12954	Enterococcus faeca
c	524	17	0.8 42324	22	AAK84724	Human immune/haema
c	525	17	0.8 55008	22	AAK80495	Human immune/haema
c	526	17	0.8 105184	24	ABK24122	Bacterial artifici
c	527	17	0.8 111309	20	AAX20250	Borrelia burgdorfe
c	528	17	0.8 128600	24	ABK83461	Human cDNA differe
	529	17	0.8 144460	21	AAZ93815	Olfactory receptor
c	530	17	0.8 160755	23	AAH88704	Human DNA sequence
	531	17	0.8 305107	22	AAH62689	Shrimp white spot
c	532	17	0.8 305107	22	AAH62689	Shrimp white spot
	533	17	0.8 349980	22	AAH68526	C glutamicum codin
c	534	17	0.8 349980	22	AAH68528	C glutamicum codin
c	535	17	0.8 349980	22	AAH68530	C glutamicum codin
c	536	17	0.8 349980	22	AAH68533	C glutamicum codin
	537	17	0.8 513445	22	AAI61373	Soybean 318O13 reg
c	538	17	0.8 513445	22	AAI61373	Soybean 318O13 reg
	539	17	0.8 580073	18	AAT58840	Mycoplasma genital
	540	17	0.8 611590	21	AAF22303	Arabidopsis thalia
	541	17	0.8 910715	20	AAX20248	Borrelia burgdorfe
c	542	17	0.8 910715	20	AAX20248	Borrelia burgdorfe
	543	17	0.8 1163020	24	ABQ67197	Listeria innocua c
	544	17	0.8 1230025	20	AAX91990	Nucleotide sequenc
c	545	17	0.8 1503900	22	AAK95240	Human neuregulin-1
c	546	17	0.8 1503900	22	AAK96733	Human neuregulin-1
c	547	17	0.8 1830121	17	AAT42063	Haemophilus influe
	548	17	0.8 2155561	24	ABN71527	Streptococcus poly
c	549	17	0.8 2155561	24	ABN71527	Streptococcus poly
c	550	17	0.8 2365589	24	ABA90521	Genomic sequence o
	551	17	0.8 3011208	24	ABQ69245	Listeria innocua D
c	552	17	0.8 4403765	22	AAI99683	Mycobacterium tube
c	553	17	0.8 4411529	22	AAI99682	Mycobacterium tube
	554	16	0.7 34	22	AAH48006	Ribosomal S9 prote
c	555	16	0.7 46	21	AAA37482	Arabidopsis thalia
	556	16	0.7 51	22	AAL26993	Human SNP oligonuc
	557	16	0.7 60	24	ABN47942	Human spliced tran
	558	16	0.7 112	22	AAK20039	Human brain expres
	559	16	0.7 112	22	AAK46083	Human bone marrow
	560	16	0.7 112	22	AAI51991	Probe #20677 used
	561	16	0.7 112	24	ABS20380	Human genome-deriv
c	562	16	0.7 118	24	ABN89706	Human ABCA10 nucle
c	563	16	0.7 119	21	AAA12629	Nucleotide sequenc
c	564	16	0.7 138	21	AAC08844	Human secreted pro
	565	16	0.7 141	22	ABA75997	Human foetal liver
	566	16	0.7 141	22	ABA40557	Probe #19023 for g
	567	16	0.7 141	22	AAK24674	Human brain expres
	568	16	0.7 141	22	AAK50673	Human bone marrow
	569	16	0.7 141	22	AAI27683	Probe #17616 for g
	570	16	0.7 141	22	AAI56651	Probe #25337 used
	571	16	0.7 141	24	ABS24145	Human genome-deriv
	572	16	0.7 148	19	AAX12107	Human biallelic po
c	573	16	0.7 155	21	AAA45873	Human secreted exp
c	574	16	0.7 155	22	ABA70749	Human foetal liver
c	575	16	0.7 155	22	ABA37250	Probe #15716 for g
c	576	16	0.7 155	22	AAK19002	Human brain expres
c	577	16	0.7 155	22	AAK44952	Human bone marrow
c	578	16	0.7 155	22	AAI25077	Probe #15010 for g
c	579	16	0.7 155	22	AAI50922	Probe #19608 used
c	580	16	0.7 155	24	ABS19196	Human genome-deriv
	581	16	0.7 178	17	AAT09783	Tomato genomic DNA
	582	16	0.7 180	24	ABL83822	Human ovarian canc
	583	16	0.7 184	21	AAC27244	Human secreted pro
c	584	16	0.7 200	16	AAT23015	Human gene signatu
c	585	16	0.7 207	22	ABA74153	Human foetal liver
c	586	16	0.7 207	22	AAK22607	Human brain expres
c	587	16	0.7 207	22	AAK48775	Human bone marrow
c	588	16	0.7 207	22	AAI54605	Probe #23291 used
c	589	16	0.7 222	20	AAX40342	Human secreted pro
c	590	16	0.7 231	24	ABN92077	Staphylococcus epi
c	591	16	0.7 234	21	AAA77922	cDNA encoding huma
c	592	16	0.7 234	22	AAI28660	Colon tumour relat
	593	16	0.7 235	23	ABV08422	Human prostate exp

c	594	16	0.7	236	23	AAS64554	DNA encoding novel
c	595	16	0.7	239	21	AAC17664	Human secreted pro
	596	16	0.7	240	21	AAC25277	Human secreted pro
c	597	16	0.7	244	21	AAC08367	Human secreted pro
c	598	16	0.7	247	22	ABA76816	Human foetal liver
c	599	16	0.7	247	22	AAK25435	Human brain expres
c	600	16	0.7	247	22	AAK51447	Human bone marrow
c	601	16	0.7	247	22	AAI57536	Probe #26222 used
c	602	16	0.7	247	24	ABS25003	Human genome-deriv
c	603	16	0.7	249	24	ABN68281	Streptococcus poly
c	604	16	0.7	250	22	AAF68109	Human lung tumour
c	605	16	0.7	250	24	ABK38020	cDNA encoding clon
	606	16	0.7	252	20	AAX51981	Human secreted pro
c	607	16	0.7	252	24	ABN26603	Human ORFX polynuc
	608	16	0.7	254	24	ABN94848	Gene #1346 used to
	609	16	0.7	254	24	ABL62614	Colon adenocarcino
	610	16	0.7	269	19	AAV59642	Human secreted pro
c	611	16	0.7	270	24	AAS61542	Lung small cell ca
c	612	16	0.7	270	24	AAS61825	Lung small cell ca
	613	16	0.7	278	16	AAT26597	Human gene signatu
	614	16	0.7	278	21	AAC01103	Human secreted pro
	615	16	0.7	278	22	AAK59778	Human immune/haema
	616	16	0.7	285	24	ABL71457	Corn tassel-derive
	617	16	0.7	290	24	ABQ58960	Human colon cancer
c	618	16	0.7	301	24	ABN26366	Human ORFX polynuc
c	619	16	0.7	304	22	ABA51322	Human breast cell
c	620	16	0.7	304	22	ABA69331	Human foetal liver
c	621	16	0.7	304	22	ABA36261	Probe #14727 for g
c	622	16	0.7	304	22	AAK17613	Human brain expres
c	623	16	0.7	304	22	AAK43424	Human bone marrow
c	624	16	0.7	304	22	AAI24204	Probe #14137 for g
c	625	16	0.7	304	22	AAI49487	Probe #18173 used
c	626	16	0.7	304	22	AAI09766	Probe #9757 used t
c	627	16	0.7	304	24	ABS17527	Human genome-deriv
	628	16	0.7	317	23	ABV31944	Human prostate exp
c	629	16	0.7	327	23	AAS50173	Staphylococcus aur
	630	16	0.7	330	22	AAI84904	Human polynucleoti
	631	16	0.7	333	23	ABV12865	Human prostate exp
c	632	16	0.7	336	15	AAQ70086	DNA encoding MAb 1
c	633	16	0.7	338	21	AAC27060	Human secreted pro
c	634	16	0.7	338	24	ABL84649	Human ovarian canc
c	635	16	0.7	346	16	AAT24869	Human gene signatu
	636	16	0.7	350	20	AAX84165	DNA encoding human
	637	16	0.7	350	21	AAC79394	5' cDNA sequence o
	638	16	0.7	350	24	ABK28938	Human breast tumou
c	639	16	0.7	351	18	AAV77280	Staphylococcus aur
c	640	16	0.7	351	22	AAH12793	Human cDNA clone (
c	641	16	0.7	353	21	AAC56123	Eucalyptus grandis
	642	16	0.7	353	22	AAK63145	Human immune/haema
c	643	16	0.7	359	23	ABV25744	Human prostate exp
	644	16	0.7	363	22	AAI80725	Human polynucleoti
	645	16	0.7	364	22	AAI80724	Human polynucleoti
c	646	16	0.7	364	22	AAS24484	Human ovarian PCR-
c	647	16	0.7	364	22	AAH83090	Human ovarian tumo
c	648	16	0.7	364	23	ABV25402	Human prostate exp
	649	16	0.7	364	23	ABV38327	Human prostate exp
c	650	16	0.7	365	24	ABN18210	Human ORFX polynuc
c	651	16	0.7	366	22	AAS24611	Human ovarian PCR-
c	652	16	0.7	366	22	AAH83233	Human ovarian tumo
c	653	16	0.7	367	21	AAC77684	Human cancer assoc
c	654	16	0.7	368	22	AAH36652	Human colon cancer
	655	16	0.7	368	23	ABV08945	Human prostate exp
	656	16	0.7	370	23	ABV49107	Human prostate exp
c	657	16	0.7	371	22	ABA59368	Human foetal liver
c	658	16	0.7	371	22	AAK07590	Human brain expres
c	659	16	0.7	371	22	AAK33388	Human bone marrow
c	660	16	0.7	371	22	AAI39172	Probe #7858 used t
c	661	16	0.7	371	24	ABS08228	Human genome-deriv
c	662	16	0.7	372	21	AAZ94513	Cartilage-associat
	663	16	0.7	373	22	AAL34973	Human musculoskele
	664	16	0.7	373	22	AAF65578	Novel human polynu
c	665	16	0.7	374	22	AAL15100	Human breast cance
	666	16	0.7	375	23	ABV10419	Human prostate exp

	667	16	0.7	377	21	AAC53208	Arabidopsis thalia
	668	16	0.7	377	22	AAI81457	Human polynucleoti
c	669	16	0.7	381	21	AAC55971	Eucalyptus grandis
c	670	16	0.7	381	21	AAC56750	Eucalyptus grandis
c	671	16	0.7	383	22	AAS38334	Novel human diagno
c	672	16	0.7	384	21	AAC56553	Eucalyptus grandis
c	673	16	0.7	386	14	AAQ60791	Human brain Expres
	674	16	0.7	387	22	AAI81458	Human polynucleoti
c	675	16	0.7	387	23	ABV13412	Human prostate exp
	676	16	0.7	389	22	AAF64371	Novel human polynu
c	677	16	0.7	389	24	ABL93110	Rat metastatic tum
c	678	16	0.7	392	22	AAF66495	Novel human polynu
	679	16	0.7	392	23	ABV38832	Human prostate exp
c	680	16	0.7	393	24	ABK80262	Bacillus clausii g
	681	16	0.7	393	24	ABL80260	Human ovarian canc
c	682	16	0.7	394	20	AAX40652	Human secreted pro
	683	16	0.7	398	21	AAC26415	Human secreted pro
c	684	16	0.7	398	23	ABV11611	Human prostate exp
	685	16	0.7	399	20	AAX40981	Human secreted pro
c	686	16	0.7	401	22	AAK95825	Human neuregulin g
c	687	16	0.7	401	22	AAK96326	Human neuregulin g
c	688	16	0.7	401	22	AAK97318	Human neuregulin g
c	689	16	0.7	401	22	AAK97819	Human neuregulin g
c	690	16	0.7	401	22	AAK74255	Human immune/haema
c	691	16	0.7	403	22	AAF66702	Novel human polynu
c	692	16	0.7	404	22	AAK61860	Human immune/haema
	693	16	0.7	409	23	ABV12151	Human prostate exp
	694	16	0.7	409	24	ABN94473	Gene #971 used to
c	695	16	0.7	410	22	AAK58995	Human immune/haema
c	696	16	0.7	425	23	ABV04243	Human prostate exp
c	697	16	0.7	427	23	ABV34530	Human prostate exp
c	698	16	0.7	427	23	ABV43388	Human prostate exp
c	699	16	0.7	429	21	AAC44637	Arabidopsis thalia
	700	16	0.7	437	23	ABV33296	Human prostate exp
	701	16	0.7	438	24	ABN18644	Human ORFX polynuc
	702	16	0.7	442	23	ABV59467	Human prostate exp
c	703	16	0.7	443	22	AAS32409	Human cDNA encodin
	704	16	0.7	443	24	ABL79528	Human ovarian canc
	705	16	0.7	446	21	AAZ80545	Human colon cancer
	706	16	0.7	458	23	ABV10608	Human prostate exp
c	707	16	0.7	459	21	AAZ94512	Cartilage-associat
	708	16	0.7	461	24	ABL80648	Human ovarian canc
c	709	16	0.7	462	20	AAZ33804	Tobacco plant resi
c	710	16	0.7	462	22	AAL23949	Human breast cance
	711	16	0.7	462	23	ABV47758	Human prostate exp
	712	16	0.7	463	23	ABV40893	Human prostate exp
c	713	16	0.7	464	21	AAC59093	Human secreted pro
c	714	16	0.7	466	21	AAC37823	Arabidopsis thalia
	715	16	0.7	467	22	ABA58772	Human foetal liver
	716	16	0.7	467	22	ABA27713	Probe #6179 for ge
	717	16	0.7	467	22	AAK06915	Human brain expres
	718	16	0.7	467	22	AAK32645	Human bone marrow
	719	16	0.7	467	22	AAI38460	Probe #7146 used t
	720	16	0.7	467	24	ABS07436	Human genome-deriv
	721	16	0.7	468	22	ABA46434	Human breast cell
	722	16	0.7	468	24	ABS05335	Human genome-deriv
	723	16	0.7	470	21	AAC03929	Human secreted pro
	724	16	0.7	471	22	AAF88291	H. tuberculata HtH
c	725	16	0.7	472	22	AAL35384	Human musculoskele
c	726	16	0.7	472	22	AAK33541	Human bone marrow
c	727	16	0.7	472	24	ABS08400	Human genome-deriv
	728	16	0.7	474	22	AAK07358	Human brain expres
	729	16	0.7	474	22	AAK33125	Human bone marrow
	730	16	0.7	474	22	AAI38921	Probe #7607 used t
	731	16	0.7	474	24	ABS07943	Human genome-deriv
c	732	16	0.7	476	18	AAV75780	Staphylococcus aur
	733	16	0.7	478	21	AAC41291	Zea mays DNA fragm
	734	16	0.7	480	24	ABN63004	Human cancer relat
	735	16	0.7	484	22	AAK54282	Human haematologic
	736	16	0.7	487	22	AAS25037	Human ovarian PCR-
c	737	16	0.7	488	21	AAC27563	Human secreted pro
	738	16	0.7	489	23	ABV10782	Human prostate exp
	739	16	0.7	492	21	AAZ43012	Human 5' EST isola

	740	16	0.7	495	24	ABN94642	Gene #1140 used to
	741	16	0.7	496	22	AAH83682	Human ovarian tumo
c	742	16	0.7	497	22	ABA57712	Human foetal liver
c	743	16	0.7	497	22	AAK05773	Human brain expres
c	744	16	0.7	497	22	AAK31397	Human bone marrow
c	745	16	0.7	497	22	AAI37290	Probe #5976 used t
c	746	16	0.7	497	24	ABS06150	Human genome-deriv
	747	16	0.7	498	22	AAH19131	Interferon alpha h
	748	16	0.7	498	22	AAH19154	Interferon alpha h
	749	16	0.7	498	22	AAH19158	Interferon alpha h
	750	16	0.7	498	22	AAH19160	Interferon alpha h
	751	16	0.7	499	24	ABL36818	Human colon tumour
c	752	16	0.7	502	22	AAK11600	Human brain expres
	753	16	0.7	505	22	AAS36292	Human cardiovascul
	754	16	0.7	510	21	AAC37182	Arabidopsis thalia
	755	16	0.7	510	24	ABQ56308	Human colon cancer
	756	16	0.7	512	23	ABV19333	Human prostate exp
c	757	16	0.7	512	23	ABV22747	Human prostate exp
c	758	16	0.7	512	23	ABV28575	Human prostate exp
c	759	16	0.7	513	24	ABN73788	Bovine embryonic g
	760	16	0.7	514	21	AAC52772	Arabidopsis thalia
c	761	16	0.7	515	24	ABN61303	Human cancer relat
c	762	16	0.7	516	22	AAS24763	Human ovarian PCR-
c	763	16	0.7	516	22	AAH83396	Human ovarian tumo
	764	16	0.7	517	24	ABK75904	Bacillus lichenifo
c	765	16	0.7	518	22	ABA61854	Human foetal liver
c	766	16	0.7	518	22	ABA29420	Probe #7886 for ge
c	767	16	0.7	518	22	AAK10167	Human brain expres
c	768	16	0.7	518	22	AAK36064	Human bone marrow
c	769	16	0.7	518	22	AAI17220	Probe #7153 for ge
c	770	16	0.7	518	22	AAI41779	Probe #10465 used
c	771	16	0.7	518	24	ABS10224	Human genome-deriv
c	772	16	0.7	522	22	AAK65802	Human immune/haema
c	773	16	0.7	523	24	ABQ52910	Oligonucleotide fo
	774	16	0.7	523	24	ABQ52911	Oligonucleotide fo
c	775	16	0.7	524	22	AAS34097	Human cDNA encodin
c	776	16	0.7	528	21	AAC75522	Human ORFX ORF1077
c	777	16	0.7	528	22	AAK61588	Human immune/haema
c	778	16	0.7	528	24	ABN21557	Human ORFX polynuc
c	779	16	0.7	530	24	ABL82245	Human ovarian canc
	780	16	0.7	533	24	ABL82317	Human ovarian canc
	781	16	0.7	540	22	AAH34721	Human colon cancer
	782	16	0.7	541	22	AAS36293	Human cardiovascul
c	783	16	0.7	544	23	ABV52539	Human prostate exp
c	784	16	0.7	547	21	AAC93806	Cat flea hindgut a
c	785	16	0.7	549	22	ABA61658	Human foetal liver
c	786	16	0.7	549	22	AAK09959	Human brain expres
c	787	16	0.7	549	22	AAK35853	Human bone marrow
c	788	16	0.7	549	22	AAI41570	Probe #10256 used
c	789	16	0.7	549	23	ABL16109	Drosophila melanog
c	790	16	0.7	550	21	AAC52560	Arabidopsis thalia
c	791	16	0.7	555	24	ABQ55631	Human ovarian anti
	792	16	0.7	565	21	AAA45987	Human metastatic m
c	793	16	0.7	566	22	ABA31368	Probe #9834 for ge
c	794	16	0.7	566	22	AAK12699	Human brain expres
c	795	16	0.7	566	22	AAK38419	Human bone marrow
c	796	16	0.7	566	22	AAI19213	Probe #9146 for ge
c	797	16	0.7	566	23	ABV54235	Human prostate exp
c	798	16	0.7	566	24	ABS12477	Human genome-deriv
	799	16	0.7	568	22	AAK61748	Human immune/haema
	800	16	0.7	570	23	ABV59365	Human prostate exp
	801	16	0.7	587	22	ABA63688	Human foetal liver
	802	16	0.7	587	22	ABA30878	Probe #9344 for ge
	803	16	0.7	587	22	AAK12204	Human brain expres
	804	16	0.7	587	22	AAK37926	Human bone marrow
	805	16	0.7	587	22	AAI18687	Probe #8620 for ge
	806	16	0.7	587	22	AAI43800	Probe #12486 used
c	807	16	0.7	587	22	AAH12019	Human cDNA clone (
c	808	16	0.7	587	23	ABV52428	Human prostate exp
	809	16	0.7	587	24	ABS11922	Human genome-deriv
c	810	16	0.7	589	22	AAS23920	Human ovarian PCR-
c	811	16	0.7	589	22	AAH82478	Human ovarian tumo
	812	16	0.7	589	23	ABV16966	Human prostate exp

c	813	16	0.7	596	21	AAF08454	Fusarium venenatum
	814	16	0.7	596	23	ABV21017	Human prostate exp
	815	16	0.7	596	23	ABV26860	Human prostate exp
	816	16	0.7	597	21	AAC87996	Human CLASP-2 geno
	817	16	0.7	597	24	ABK84988	DNA encoding cadhe
	818	16	0.7	606	21	AAC34869	Arabidopsis thalia
c	819	16	0.7	607	21	AAF11684	Aspergillus niger
c	820	16	0.7	607	24	ABN65920	Human cancer relat
c	821	16	0.7	608	23	ABV55709	Human prostate exp
	822	16	0.7	609	23	ABL16393	Drosophila melanog
c	823	16	0.7	611	24	ABQ59289	Human colon cancer
	824	16	0.7	612	21	AAC34957	Arabidopsis thalia
	825	16	0.7	614	21	AAC03943	Human secreted pro
c	826	16	0.7	615	22	ABA12283	Human nervous syst
	827	16	0.7	616	21	AAC54217	Arabidopsis thalia
c	828	16	0.7	617	21	AAC52556	Arabidopsis thalia
	829	16	0.7	619	22	AAL22899	Human breast cance
c	830	16	0.7	624	23	ABL21093	Drosophila melanog
	831	16	0.7	624	24	ABQ90302	M. capsulatus gene
	832	16	0.7	625	24	ABQ66171	Arabidopsis thalia
	833	16	0.7	626	21	AAA64650	Partial sequence T
	834	16	0.7	626	23	ABV52604	Human prostate exp
c	835	16	0.7	640	24	ABQ68279	Listeria monocytog
c	836	16	0.7	640	24	ABQ70187	Listeria monocytog
	837	16	0.7	641	19	AAV52492	Streptococcus pneu
c	838	16	0.7	652	21	AAC76125	Human ORFX ORF1680
	839	16	0.7	654	21	AAC34319	Arabidopsis thalia
c	840	16	0.7	655	21	AAF14864	Aspergillus oryzae
	841	16	0.7	659	19	AAX14058	H. pylori GHPO 257
c	842	16	0.7	661	24	ABK35446	Human cDNA encodin
	843	16	0.7	662	23	AAS75511	DNA encoding novel
c	844	16	0.7	667	21	AAC50291	Arabidopsis thalia
	845	16	0.7	670	21	AAC77246	Human ORFX ORF2801
c	846	16	0.7	672	21	AAC33023	Arabidopsis thalia
	847	16	0.7	674	24	ABN79034	Human ORF3981 cDNA
c	848	16	0.7	687	21	AAF12344	Aspergillus oryzae
	849	16	0.7	687	22	AAH53771	S. epidermidis ope
	850	16	0.7	687	22	AAH53901	S. epidermidis ope
	851	16	0.7	688	23	ABV46761	Human prostate exp
	852	16	0.7	690	24	ABK27480	DNA encoding novel
	853	16	0.7	694	23	AAS52194	Staphylococcus aur
	854	16	0.7	698	24	ABQ71089	Listeria monocytog
c	855	16	0.7	702	22	ABA18214	Human nervous syst
c	856	16	0.7	702	22	AAS36731	Human cardiovascul
c	857	16	0.7	702	22	AAS36732	Human cardiovascul
	858	16	0.7	702	24	ABN91343	Staphylococcus epi
c	859	16	0.7	704	22	ABA18216	Human nervous syst
c	860	16	0.7	704	22	AAS36733	Human cardiovascul
	861	16	0.7	708	22	ABA13504	Human nervous syst
c	862	16	0.7	714	20	AAX37430	Human secreted pro
	863	16	0.7	732	22	AAI96975	Human neuroblastom
	864	16	0.7	739	21	AAC10885	Human secreted pro
	865	16	0.7	740	20	AAZ16154	Human gene express
	866	16	0.7	740	20	AAZ15294	Human gene express
	867	16	0.7	750	21	AAA79441	Eucalyptus.grandis
c	868	16	0.7	755	21	AAZ64946	Membrane-bound pro
c	869	16	0.7	755	22	AAF44092	Human EST DNA14232
c	870	16	0.7	759	24	ABN99200	Arabidopsis thalia
c	871	16	0.7	760	20	AAX39750	Gastric cancer ass
c	872	16	0.7	760	22	AAK57151	Human immune/haema
	873	16	0.7	762	23	AAS51797	Staphylococcus aur
	874	16	0.7	765	23	AAS54539	Staphylococcus aur
	875	16	0.7	766	22	AAL10335	Human breast cance
c	876	16	0.7	767	21	AAF09451	Fusarium venenatum
	877	16	0.7	772	22	AAL19915	Human breast cance
c	878	16	0.7	774	22	AAK62700	Human immune/haema
	879	16	0.7	776	22	AAH08104	Human cDNA clone (
	880	16	0.7	781	22	AAH07611	Human cDNA clone (
	881	16	0.7	782	19	AAV59777	Human secreted pro
	882	16	0.7	787	22	AAH05040	Human cDNA clone (
c	883	16	0.7	801	21	AAC46622	Zea mays DNA fragm
	884	16	0.7	804	24	ABN99034	Arabidopsis thalia
	885	16	0.7	804	24	ABN66101	Streptococcus poly
	886	16	0.7	804	24	ABN70282	Streptococcus poly
	887	16	0.7	805	22	AAH05496	Human cDNA clone (
	888	16	0.7	815	23	AAS64555	DNA encoding novel
	889	16	0.7	824	24	ABQ43686	Oligonucleotide fo
c	890	16	0.7	824	24	ABQ43687	Oligonucleotide fo
c	891	16	0.7	826	22	AAH06926	Human cDNA clone (
c	892	16	0.7	829	22	AAH72786	Human cervical can
c	893	16	0.7	833	21	AAC75348	Human ORFX ORF903
	894	16	0.7	833	24	ABN98918	Arabidopsis thalia
c	895	16	0.7	833	24	ABN78409	Human kinase-like
	896	16	0.7	836	21	AAA64655	Partial sequence T
	897	16	0.7	837	24	ABN68887	Streptococcus poly
c	898	16	0.7	841	20	AAX36815	Human transmembran
c	899	16	0.7	841	23	ABL23981	Drosophila melanog
	900	16	0.7	843	23	ABL25963	Drosophila melanog
c	901	16	0.7	846	22	AAL22796	Human breast cance
	902	16	0.7	846	24	ABT04077	Human ovary specif
	903	16	0.7	847	22	AAK93797	Human cDNA clone r
	904	16	0.7	847	22	AAH08614	Human cDNA clone (
	905	16	0.7	855	24	ABN66088	Streptococcus poly
c	906	16	0.7	856	21	AAC43619	Arabidopsis thalia
c	907	16	0.7	858	22	AAH05559	Human cDNA clone (
	908	16	0.7	858	23	AAS72767	DNA encoding novel
c	909	16	0.7	860	20	AAX97685	Extended human sec
c	910	16	0.7	870	20	AAZ33432	Human prostate can
	911	16	0.7	870	24	ABK79215	Bacillus clausii g
c	912	16	0.7	879	20	AAX39743	Gastric cancer ass
	913	16	0.7	883	24	ABQ39734	Oligonucleotide fo
c	914	16	0.7	883	24	ABQ39735	Oligonucleotide fo
c	915	16	0.7	891	21	AAA87704	Human secreted pro
c	916	16	0.7	903	21	AAC51039	Arabidopsis thalia
c	917	16	0.7	906	22	AAS46195	Human DNA encoding
c	918	16	0.7	917	24	ABQ23332	Oligonucleotide fo
	919	16	0.7	917	24	ABQ23333	Oligonucleotide fo
	920	16	0.7	934	17	AAT05843	Repeat sequence of
	921	16	0.7	934	20	AAZ11801	Repeat upstream fl
	922	16	0.7	934	22	AAK61467	Human immune/haema
c	923	16	0.7	935	20	AAZ25017	Human channel-rela
	924	16	0.7	937	22	ABA16235	Human nervous syst
c	925	16	0.7	937	22	ABA16235	Human nervous syst
	926	16	0.7	937	22	ABA16236	Human nervous syst
c	927	16	0.7	937	22	ABA16236	Human nervous syst
	928	16	0.7	937	22	ABA16237	Human nervous syst
c	929	16	0.7	937	22	ABA16237	Human nervous syst
c	930	16	0.7	937	22	ABA09289	Human channel-rela
c	931	16	0.7	937	22	AAH31714	Human olfactory re
	932	16	0.7	940	22	AAI59761	Human polynucleoti
c	933	16	0.7	940	22	AAD05633	Human secreted pro
c	934	16	0.7	941	20	AAV84586	Human secreted pro
c	935	16	0.7	941	21	AAA87708	Human secreted pro
c	936	16	0.7	941	22	ABA83369	Human secreted pro
c	937	16	0.7	941	24	AAL45808	Human secreted pro
c	938	16	0.7	956	21	AAC50116	Arabidopsis thalia
c	939	16	0.7	956	22	AAI57975	Human polynucleoti
c	940	16	0.7	958	21	AAA87706	Human secreted pro
c	941	16	0.7	960	21	AAC33700	Arabidopsis thalia
c	942	16	0.7	960	24	ABQ67738	Listeria innocua D
c	943	16	0.7	960	24	ABQ69345	Listeria innocua D
c	944	16	0.7	960	24	ABQ69700	Listeria innocua D
c	945	16	0.7	962	23	ABV25053	Human prostate exp
	946	16	0.7	972	24	ABK71683	Human dithp polynu
c	947	16	0.7	976	22	AAH77112	DNA sequence encod
	948	16	0.7	981	22	AAK80202	Human immune/haema
c	949	16	0.7	988	24	AAL45833	Human secreted pro
	950	16	0.7	1010	20	AAX84933	Human secreted pro
c	951	16	0.7	1011	22	AAH52615	S. epidermidis ope
c	952	16	0.7	1011	22	AAH53603	S. epidermidis ope
c	953	16	0.7	1020	21	AAA87707	Human secreted pro
c	954	16	0.7	1044	22	AAH72613	Human cervical can
c	955	16	0.7	1052	24	ABL90766	Human polynucleoti
c	956	16	0.7	1053	24	AAK98792	Loblolly pine cDNA
	957	16	0.7	1056	23	ABL26027	Drosophila melanog
c	958	16	0.7	1060	22	AAF88393	M. crenulata KLH2

```
  959     16    0.7   1068  21  AAC33750            Arabidopsis thalia
  960     16    0.7   1071  21  AAC42723            Arabidopsis thalia
  961     16    0.7   1074  21  AAC50451            Arabidopsis thalia
c 962     16    0.7   1078  21  AAC46146            Arabidopsis thalia
c 963     16    0.7   1080  21  AAC33643            Arabidopsis thalia
  964     16    0.7   1080  22  AAH99353            Human protein enco
c 965     16    0.7   1080  22  AAH99353            Human protein enco
  966     16    0.7   1082  21  AAC39766            Arabidopsis thalia
  967     16    0.7   1104  15  AAQ72616            Insecticidal patat
  968     16    0.7   1110  21  AAC35938            Arabidopsis thalia
c 969     16    0.7   1110  22  AAF27709            Human transport pr
  970     16    0.7   1111  21  AAC32644            Arabidopsis thalia
c 971     16    0.7   1115  23  ABV20220            Human prostate exp
c 972     16    0.7   1115  23  ABV20247            Human prostate exp
c 973     16    0.7   1115  23  ABV20287            Human prostate exp
c 974     16    0.7   1115  23  ABV26049            Human prostate exp
c 975     16    0.7   1115  23  ABV26078            Human prostate exp
c 976     16    0.7   1115  23  ABV26118            Human prostate exp
  977     16    0.7   1120  22  AAL01967            Human reproductive
  978     16    0.7   1120  23  ABL97260            Human testicular a
c 979     16    0.7   1121  22  AAH98814            Human EST-derived
  980     16    0.7   1124  18  AAT74083            Murine LS-DNase nu
c 981     16    0.7   1125  20  AAX13428            Enterococcus faeca
c 982     16    0.7   1130  21  AAA87693            Human secreted pro
  983     16    0.7   1136  19  AAV37396            Streptococcus pneu
c 984     16    0.7   1139  21  AAC33237            Arabidopsis thalia
  985     16    0.7   1143  22  AAL07044            Human reproductive
  986     16    0.7   1180  18  AAX30797            Streptococcus pneu
c 987     16    0.7   1181  19  AAT98658            DNA encoding a S.
  988     16    0.7   1192  20  AAX37262            Rat heparanase enz
  989     16    0.7   1203  21  AAC43685            Zea mays DNA fragm
c 990     16    0.7   1205  21  AAC47687            Zea mays DNA fragm
  991     16    0.7   1205  23  ABL10657            Drosophila melanog
  992     16    0.7   1208  22  AAH98949            Human EST-derived
  993     16    0.7   1224  22  ABA19704            Human nervous syst
c 994     16    0.7   1231  24  ABK65217            Arabidopsis cDNA e
  995     16    0.7   1238  22  ABA13121            Human nervous syst
c 996     16    0.7   1242  23  AAS75527            DNA encoding novel
c 997     16    0.7   1249  19  AAV38987            cDNA encoding a hu
c 998     16    0.7   1249  21  AAA58814            DNA encoding a hum
c 999     16    0.7   1251  21  AAA98916            H. tuberculata hem
c1000     16    0.7   1251  21  AAA98944            M. crenulata hemoc
```

ALIGNMENTS

```
RESULT 1
AAD06354
ID   AAD06354 standard; cDNA; 2224 BP.
XX
AC   AAD06354;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human vasa cDNA.
XX
KW   Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;
KW   ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;
KW   colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;
KW   skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;
KW   medulloblastoma; choriocarcinoma; squamous cell carcinoma; leukaemia;
KW   acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;
KW   osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KW   leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW   fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KW   teratoma; mediastinal; intracranial; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             14..2188
FT                   /*tag=  a
FT                   /product= "Human vasa protein"
FT   misc_feature    14..2185
FT                   /*tag=  b
FT                   /note= "This region is specifically claimed as
FT                   SEQ ID NO: 15 in claim 3 of the specification"
XX
PN   WO200136445-A1.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31485.
XX
PR   18-NOV-1999;   99US-0166394.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Castrillon DH;
XX
DR   WPI; 2001-355606/37.
DR   P-PSDB; AAE02417.
XX
PT   Novel vasa polynucleotides useful in the diagnosis or treatment of
PT   conditions characterized by aberrant expression and/or presence of
PT   mutant forms of vasa polynucleotides or polypeptides -
XX
PS   Claim 1; Page 52-53; 66pp; English.
XX
CC   The present sequence is human vasa cDNA that has germ cell specific
CC   expression and is believed to play a determinative role in gonad
CC   development. Germ cells are specialised to produce haploid gametes in
CC   multicellular organisms. Vasa is useful in the diagnosis or treatment
CC   of conditions characterised by its aberrant expression and/or the
CC   presence of its mutant forms. The conditions include cancers such as
CC   biliary tract, brain, breast, colon, ovarian, pancreas, prostate,
CC   colorectal, oral, liver, lung, skin, basocellular, testis, renal,
CC   thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,
CC   melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC   choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC   acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC   Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC   intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC   such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC   Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC   (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC   of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
SQ   Sequence 2224 BP; 678 A; 385 C; 550 G; 611 T; 0 other;

  Query Match             100.0%;  Score 2172;  DB 22;  Length 2224;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2172;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 73

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 133

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 193

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 253

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 313

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          314 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 373

Qy          361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          374 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 433

Qy          421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          434 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 493

Qy          481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          494 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 553

Qy          541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          554 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 613

Qy          601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          614 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 673

Qy          661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          674 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 733

Qy          721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          734 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 793

Qy          781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          794 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 853

Qy          841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          854 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 913

Qy          901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          914 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 973

Qy          961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          974 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1033

Qy         1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1034 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1093

Qy         1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1094 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1153

Qy         1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1154 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1213

Qy         1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1214 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1273

Qy         1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1274 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1333

Qy         1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1334 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1393

Qy         1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1394 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1453

Qy         1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1454 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1513

Qy         1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1514 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1573

Qy         1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1574 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1633

Qy         1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1634 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1693

Qy         1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1694 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1753

Qy         1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1754 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1813

Qy         1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1814 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1873

Qy         1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1874 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1933

Qy         1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1934 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1993

Qy         1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1994 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2053

Qy         2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2054 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2113

Qy         2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2114 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2173

Qy         2161 GAGTCATGGGAT 2172
                ||||||||||||
Db         2174 GAGTCATGGGAT 2185


RESULT 2
AAA44726
ID   AAA44726 standard; cDNA; 1006 BP.
XX
AC   AAA44726;
XX
DT   21-AUG-2000  (first entry)
XX
DE   Human secreted expressed sequence tag SEQ ID NO:1301.
XX
KW   Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
KW   expressed sequence tag; EST; probe; chemotactic; proliferative;
KW   immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW   thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW   antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW   antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW   cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW   vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
```

```
KW   insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW   lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW   central nervous system disorder; Alzheimer's disease; stroke;
KW   Parkinson's disease; Huntington's disease; coagulation disorder;
KW   haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW   tumour; infection; depression; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200021991-A1.
XX
PD   20-APR-2000.
XX
PF   15-OCT-1999;   99WO-US24206.
XX
PR   15-OCT-1998;   98US-0104436.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  McCoy JM,  LaVallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M,  Bowman MR;
XX
DR   WPI; 2000-317938/27.
XX
PT   Isolated polynucleotides, and encoded proteins, comprising secreted
PT   expressed sequence tags (sESTs), useful for treating various disorders
PT   such as autoimmune, infectious, and central nervous system disorders -
XX
PS   Claim 1; Page 540; 803pp; English.
XX
CC   AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC   sequence tags (sESTs), isolated from human, mouse, chicken and rat
CC   tissue sources. The sESTs can have a range of activities depending on
CC   the tissues they were isolated from. The activities include:
CC   chemotactic; proliferative; immunomodulatory; haematopoietic;
CC   chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC   cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC   antiasthmatic; vulnerary; antiulcer; osteopathic; neuroprotective;
CC   nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC   anticonvulsant; and antidepressant. The sESTs can be used for gene
CC   therapy and in vaccines. The sESTs are useful as probes for the
CC   identification and isolation of full-length cDNAs and genomic DNA
CC   molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC   are useful in assays for determining biological activity and raising
CC   antibodies. They may be useful for treatment of autoimmune disorders
CC   (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC   (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC   osteoporosis, osteoarthritis, central nervous system disorders
CC   (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC   disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC   disease), tumours, bacterial, fungal or viral infections, depression and
CC   psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC   in the exemplification of the present invention.
XX
SQ   Sequence 1006 BP; 314 A; 181 C; 251 G; 260 T; 0 other;

  Query Match            26.7%;  Score 579;  DB 21;  Length 1006;
  Best Local Similarity  99.8%;  Pred. No. 5.6e-276;
  Matches  629;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        497 ATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAA 436

Qy        557 GACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAA 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAA 496

Qy        617 GTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATT 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 GTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATT 556

Qy        677 CTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 CTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCT 616

Qy        737 ACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCA 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 ACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCA 676

Qy        797 TAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAG 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 TAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAG 736

Qy        857 CAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAG 916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 CAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAG 796

Qy        917 CTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGAC 976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 CTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGAC 856

Qy        977 GAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAA 1036
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        857 GAGATTTGATGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAA 916

Qy       1037 TTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAAC 1096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 TTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAAC 976

Qy       1097 CAGAGTGTATTATTGTAGCACCAACTCGAG 1126
              ||||||||||||||||||||||||||||||
Db        977 CAGAGTGTATTATTGTAGCACCAACTCGAG 1006


RESULT 3
AAS40979
ID   AAS40979 standard; cDNA; 677 BP.
XX
AC   AAS40979;
XX
DT   17-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human enzyme polypeptide #195.
XX
KW   Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW   ligase; hyperproliferative disorder; immunodeficiency disorder;
KW   autoimmune disorder; neurological disorder; metabolic disorder;
KW   inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW   blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW   anti arthritic; nephrotropic; anticoagulant; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155301-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01239.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
```

```
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465566/50.
DR   P-PSDB; AAU23109.
XX
PT   Novel polypeptides and polynucleotides useful for diagnosing,
PT   preventing, treating neural, immune system, muscular, reproductive,
PT   pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT   diseases  -
XX
PS   Claim 4; SEQ ID No 205; 1180pp; English.
XX
CC   The present invention relates to the isolation of novel human enzyme
CC   polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC   encoding them. The enzyme polypeptides of the invention may comprise the
CC   functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC   isomerases or ligases. The sequences of the invention are useful in the
CC   diagnosis, treatment, prevention and/or prognosis of a wide range of
CC   disorders including hyperproliferative disorders (e.g. cancer),
CC   immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC   (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC   metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC   (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
```

```
CC   blood-related disorders (e.g. haemophilia), reproductive disorders
CC   (e.g. infertility) and infectious disorders (e.g. Influenza). The
CC   polynucleotides of the invention can also be used in gene therapy.
CC   AAS40785-AAS41684 represent cDNA sequences encoding for the novel human
CC   enzyme polypeptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 677 BP; 225 A; 117 C; 130 G; 205 T; 0 other;

  Query Match             17.0%;  Score 370;  DB 22;  Length 677;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-172;
  Matches  370;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 240

Qy       2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 300

Qy       2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 360

Qy       2128 TCTTCTTCAC 2137
              ||||||||||
Db        361 TCTTCTTCAC 370


RESULT 4
AAS41534
ID   AAS41534 standard; cDNA; 591 BP.
XX
AC   AAS41534;
XX
DT   17-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human enzyme polypeptide #750.
XX
KW   Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW   ligase; hyperproliferative disorder; immunodeficiency disorder;
KW   autoimmune disorder; neurological disorder; metabolic disorder;
KW   inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW   blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW   anti arthritic; nephrotropic; anticoagulant; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155301-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01239.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
```

```
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465566/50.
DR   P-PSDB; AAU23664.
XX
PT   Novel polypeptides and polynucleotides useful for diagnosing,
PT   preventing, treating neural, immune system, muscular, reproductive,
PT   pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT   diseases  -
XX
PS   Claim 4; SEQ ID No 760; 1180pp; English.
XX
CC   The present invention relates to the isolation of novel human enzyme
CC   polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC   encoding them. The enzyme polypeptides of the invention may comprise the
CC   functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC   isomerases or ligases. The sequences of the invention are useful in the
CC   diagnosis, treatment, prevention and/or prognosis of a wide range of
CC   disorders including hyperproliferative disorders (e.g. cancer),
CC   immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC   (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC   metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC   (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC   blood-related disorders (e.g. haemophilia), reproductive disorders
CC   (e.g. infertility) and infectious disorders (e.g. Influenza). The
CC   polynucleotides of the invention can also be used in gene therapy.
CC   AAS40785-AAS41684 represent cDNA sequences encoding for the novel human
CC   enzyme polypeptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 591 BP; 203 A; 100 C; 111 G; 176 T; 1 other;

  Query Match             11.7%;  Score 255;  DB 22;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-115;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1883 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAAT 1942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAAT 89

Qy       1943 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 2002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 149

Qy       2003 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAA 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAA 209

Qy       2063 GAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG 269

Qy       2123 GGTTTTCTTCTTCAC 2137
              |||||||||||||||
Db        270 GGTTTTCTTCTTCAC 284


RESULT 5
AAL01169
ID   AAL01169 standard; cDNA; 591 BP.
XX
AC   AAL01169;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen cDNA SEQ ID NO: 1170.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
```

```
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465570/50.
DR   P-PSDB; AAM95199.
XX
PT   Isolated nucleic acid molecule encoding a reproductive system antigen
PT   is used in preventing, treating or ameliorating a medical condition  -
XX
PS   Claim 1; SEQ ID NO 1170; 1297pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
```

```
CC   number of human reproductive system related antigens. These can be used
CC   in the prevention and treatment of reproductive system disorders,
CC   including cancer. The present sequence is a coding sequence of the
CC   invention.
XX
SQ   Sequence 591 BP; 203 A; 100 C; 111 G; 176 T; 1 other;

  Query Match             11.7%;  Score 255;  DB 22;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-115;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1883 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAAT 1942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAAT 89

Qy       1943 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 2002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 149

Qy       2003 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAA 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAA 209

Qy       2063 GAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG 269

Qy       2123 GGTTTTCTTCTTCAC 2137
              |||||||||||||||
Db        270 GGTTTTCTTCTTCAC 284


RESULT 6
ABL96630
ID   ABL96630 standard; cDNA; 591 BP.
XX
AC   ABL96630;
XX
DT   21-JUN-2002  (first entry)
XX
DE   Human testicular antigen encoding cDNA SEQ ID NO: 298.
XX
KW   Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW   reproductive system disorder; urinary system disorder; gene therapy;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disease; infection; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155317-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01329.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
```

```
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483232/52.
XX
PT   Nucleic acids encoding 973 human testicular antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating testicular cancer  -
XX
PS   Claim 1; SEQ ID NO 298; 766pp; English.
XX
CC   The present invention provides the protein and coding sequences of 973
CC   human testicular antigens, and fragments of their genomic sequences. The
CC   sequences can be used in the treatment of cardiovascular, urinary system,
CC   reproductive system, immune, respiratory, neurological and
CC   gastrointestinal disorders, infections, and particularly cancer,
CC   especially testicular cancers. The present sequence is a cDNA of the
CC   invention.
XX
SQ   Sequence 591 BP; 203 A; 100 C; 111 G; 176 T; 1 other;

  Query Match             11.7%;  Score 255;  DB 23;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-115;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1883 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAAT 1942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAAT 89

Qy       1943 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 2002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 149

Qy       2003 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAA 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAA 209

Qy       2063 GAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTG 269

Qy       2123 GGTTTTCTTCTTCAC 2137
              |||||||||||||||
Db        270 GGTTTTCTTCTTCAC 284


RESULT 7
AAL04793
ID   AAL04793 standard; DNA; 2194 BP.
XX
AC   AAL04793;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen DNA SEQ ID NO: 7481.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
```

```
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465570/50.
XX
PT   Isolated nucleic acid molecule encoding a reproductive system antigen
PT   is used in preventing, treating or ameliorating a medical condition  -
XX
PS   Disclosure; SEQ ID NO 7481; 1297pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human reproductive system related antigens. These can be used
CC   in the prevention and treatment of reproductive system disorders,
CC   including cancer. The present sequence is a genomic sequence encoding a
CC   protein of the invention.
XX
SQ   Sequence 2194 BP; 624 A; 349 C; 400 G; 821 T; 0 other;

  Query Match             10.1%;  Score 220;  DB 22;  Length 2194;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-98;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              ||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 8
ABL97687
```

```
ID   ABL97687 standard; DNA; 2194 BP.
XX
AC   ABL97687;
XX
DT   21-JUN-2002  (first entry)
XX
DE   Human testicular antigen encoding DNA fragment SEQ ID NO: 2339.
XX
KW   Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW   reproductive system disorder; urinary system disorder; gene therapy;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155317-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01329.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
```

```
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483232/52.
XX
PT   Nucleic acids encoding 973 human testicular antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating testicular cancer  -
XX
PS   Disclosure; SEQ ID NO 2339; 766pp; English.
XX
CC   The present invention provides the protein and coding sequences of 973
CC   human testicular antigens, and fragments of their genomic sequences. The
CC   sequences can be used in the treatment of cardiovascular, urinary system,
CC   reproductive system, immune, respiratory, neurological and
CC   gastrointestinal disorders, infections, and particularly cancer,
CC   especially testicular cancers. The present sequence is a DNA encoding a
CC   protein fragment of the invention.
XX
SQ   Sequence 2194 BP; 624 A; 349 C; 400 G; 821 T; 0 other;

  Query Match             10.1%;  Score 220;  DB 23;  Length 2194;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-98;
  Matches  220; Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              ||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 9
AAL04794
ID   AAL04794 standard; DNA; 2195 BP.
XX
AC   AAL04794;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen DNA SEQ ID NO: 7482.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
```

```
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465570/50.
XX
PT   Isolated nucleic acid molecule encoding a reproductive system antigen
PT   is used in preventing, treating or ameliorating a medical condition  -
XX
PS   Disclosure; SEQ ID NO 7482; 1297pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human reproductive system related antigens. These can be used
CC   in the prevention and treatment of reproductive system disorders,
CC   including cancer. The present sequence is a genomic sequence encoding a
CC   protein of the invention.
XX
SQ   Sequence 2195 BP; 624 A; 350 C; 400 G; 821 T; 0 other;

  Query Match             10.1%;  Score 220;  DB 22;  Length 2195;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-98;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              ||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 10
ABL97688
ID   ABL97688 standard; DNA; 2195 BP.
XX
AC   ABL97688;
XX
DT   21-JUN-2002  (first entry)
XX
DE   Human testicular antigen encoding DNA fragment SEQ ID NO: 2340.
XX
KW   Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW   reproductive system disorder; urinary system disorder; gene therapy;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155317-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01329.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
```

```
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483232/52.
XX
PT   Nucleic acids encoding 973 human testicular antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating testicular cancer  -
XX
PS   Disclosure; SEQ ID NO 2340; 766pp; English.
XX
CC   The present invention provides the protein and coding sequences of 973
CC   human testicular antigens, and fragments of their genomic sequences. The
CC   sequences can be used in the treatment of cardiovascular, urinary system,
CC   reproductive system, immune, respiratory, neurological and
CC   gastrointestinal disorders, infections, and particularly cancer,
CC   especially testicular cancers. The present sequence is a DNA encoding a
CC   protein fragment of the invention.
XX
SQ   Sequence 2195 BP; 624 A; 350 C; 400 G; 821 T; 0 other;

  Query Match             10.1%;  Score 220;  DB 23;  Length 2195;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-98;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy      1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy      1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy      1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
             ||||||||||||||||||||||||||||||||||||||||
Db       181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 11
AAL04795
ID   AAL04795 standard; DNA; 2197 BP.
XX
AC   AAL04795;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen DNA SEQ ID NO: 7483.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
```

```
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465570/50.
XX
PT   Isolated nucleic acid molecule encoding a reproductive system antigen
PT   is used in preventing, treating or ameliorating a medical condition  -
XX
PS   Disclosure; SEQ ID NO 7483; 1297pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human reproductive system related antigens. These can be used
CC   in the prevention and treatment of reproductive system disorders,
CC   including cancer. The present sequence is a genomic sequence encoding a
CC   protein of the invention.
XX
SQ   Sequence 2197 BP; 624 A; 349 C; 401 G; 823 T; 0 other;

  Query Match             10.1%;  Score 220;  DB 22;  Length 2197;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-98;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              ||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 12
ABL97689
ID   ABL97689 standard; DNA; 2197 BP.
XX
AC   ABL97689;
XX
DT   21-JUN-2002  (first entry)
XX
DE   Human testicular antigen encoding DNA fragment SEQ ID NO: 2341.
XX
KW   Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW   reproductive system disorder; urinary system disorder; gene therapy;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155317-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01329.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
```

```
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483232/52.
XX
PT   Nucleic acids encoding 973 human testicular antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating testicular cancer  -
XX
PS   Disclosure; SEQ ID NO 2341; 766pp; English.
XX
CC   The present invention provides the protein and coding sequences of 973
CC   human testicular antigens, and fragments of their genomic sequences. The
CC   sequences can be used in the treatment of cardiovascular, urinary system,
CC   reproductive system, immune, respiratory, neurological and
CC   gastrointestinal disorders, infections, and particularly cancer,
CC   especially testicular cancers. The present sequence is a DNA encoding a
CC   protein fragment of the invention.
XX
SQ   Sequence 2197 BP; 624 A; 349 C; 401 G; 823 T; 0 other;

  Query Match             10.1%;  Score 220;  DB 23;  Length 2197;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-98;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              ||||||||||||||||||||||||||||||||||||||||
Db        181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 13
ABN42617
ID   ABN42617 standard; DNA; 60 BP.
XX
AC   ABN42617;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:15365.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
```

```
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 15365; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 19 A; 11 C; 18 G; 12 T; 0 other;

  Query Match             2.8%;  Score 60;  DB 24;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-19;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGG 2124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGG 60

RESULT 14
ABI99243
ID   ABI99243 standard; cDNA; 2187 BP.
XX
AC   ABI99243;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Mouse ischaemic condition related cDNA sequence SEQ ID NO:67.
XX
KW   Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW   vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS   Mus musculus.
XX
PN   WO200188188-A2.
XX
PD   22-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-JP04192.
XX
PR   18-MAY-2000; 2000JP-0145977.
XX
PA   (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI   Ishikawa K,  Asai S,  Takahashi Y,  Nagata T,  Ishii Y;
XX
DR   WPI; 2002-034733/04.
DR   P-PSDB; ABB57045.
XX
PT   Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT   expression levels of particular genes defined in the specification or
PT   by determining the expression profile of a gene group comprising these
PT   genes -
XX
PS   Claim 2; Page 210-215; 2690pp; English.
XX
CC   The present invention describes a method for examining ischaemic
CC   conditions, comprising measuring the expression levels of particular
CC   genes (I) in a test sample or determining the expression profile of a
CC   gene group in the sample comprising genes selected from (I). The method
CC   is useful for examining the ischaemic condition (e.g. compressive
CC   ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC   expression levels of particular genes (ABI99202 to ABI99912, encoding
CC   the protein sequences in ABB57020 to ABB57374) or by determining the
CC   expression profile of a gene group comprising these genes. The
CC   expression levels or expression profiles produced by these genes are
CC   used as an indicator when screening for ischaemic condition-improving
CC   drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC   represent PCR primers for a mouse ischaemic condition related sequence,
CC   which are used in the exemplification of the present invention.
XX
SQ   Sequence 2187 BP; 635 A; 409 C; 596 G; 547 T; 0 other;

  Query Match             1.1%;  Score 23;  DB 24;  Length 2187;
  Best Local Similarity   100.0%;  Pred. No. 0.83;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        982 TTGATGGCTTGTGCTCAAACAGG 1004
              |||||||||||||||||||||||
Db        681 TTGATGGCTTGTGCTCAAACAGG 703

RESULT 15
AAS53256
ID   AAS53256 standard; DNA; 1842 BP.
XX
AC   AAS53256;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Haemophilus influenzae DNA for cellular proliferation protein #38.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Haemophilus influenzae.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
```

```
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU35397.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 6893; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1842 BP; 560 A; 381 C; 392 G; 509 T; 0 other;

  Query Match             1.0%;  Score 21;  DB 23;  Length 1842;
  Best Local Similarity   100.0%;  Pred. No. 8.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1879 CGAATTGGGCGTACTGGTCGT 1899
              |||||||||||||||||||||
Db        991 CGAATTGGGCGTACTGGTCGT 1011


Search completed: June 10, 2003, 12:33:38
Job time : 526 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:          June 10, 2003, 10:42:49 ; Search time 5458 Seconds
                                           (without alignments)
                                           11581.402 Million cell updates/sec

Title:           US-09-714-865-15
Perfect score:   2172
Sequence:        1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:   OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0

Searched:        2054640 seqs, 14551402878 residues

Word size :       0

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :        GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2172	100.0	2224	9	AY004154	AY004154 Homo sapi
	2	1978	91.1	2411	9	AF262962	AF262962 Homo sapi
	3	1719	79.1	2189	9	HSM802178	AL137462 Homo sapi
	4	1265	58.2	1984	9	AK093439	AK093439 Homo sapi
	5	273	12.6	139677	9	AC008914	AC008914 Homo sapi
c	6	273	12.6	176784	2	AC016632	AC016632 Homo sapi
c	7	273	12.6	182126	2	AC016639	AC016639 Homo sapi
	8	273	12.6	204250	2	AC022265	AC022265 Homo sapi
c	9	174	8.0	204250	2	AC022265	AC022265 Homo sapi
	10	58	2.7	524	4	AY100475	AY100475 Equus cab
	11	53	2.4	3030	10	S75275	S75275 RVLG=vasa-l
	12	50	2.3	1930	10	MUSDVH	D14859 Mouse mRNA
c	13	47	2.2	108795	2	AC124894	AC124894 Rattus no
c	14	47	2.2	114024	2	AC105884	AC105884 Rattus no
c	15	36	1.7	14350	1	AE013856	AE013856 Yersinia
	16	36	1.7	335050	1	AJ414150	AJ414150 Yersinia
	17	31	1.4	62682	2	AC109144	AC109144 Mus muscu
	18	31	1.4	205057	2	AC108418	AC108418 Mus muscu
	19	29	1.3	192541	2	AC127770	AC127770 Rattus no
	20	28	1.3	2721	3	AB047382	AB047382 Hydra mag
	21	27	1.2	108795	2	AC124894	AC124894 Rattus no
	22	26	1.2	3731	10	MMU7376	AJ007376 Mus muscu
	23	26	1.2	3767	10	BC021453	BC021453 Mus muscu
c	24	26	1.2	62682	2	AC109144	AC109144 Mus muscu
	25	26	1.2	114024	2	AC105884	AC105884 Rattus no
c	26	25	1.2	178133	2	AC123925	AC123925 Mus muscu
	27	23	1.1	1989	5	AB004836	AB004836 Gallus ga
	28	23	1.1	2187	6	AX305316	AX305316 Sequence
	29	23	1.1	2187	10	MUSRNAHELI	L25126 Mus musculu
	30	23	1.1	3132	3	AB047383	AB047383 Hydra mag
	31	23	1.1	3185	10	MMDBRNAHL	Z38117 M.musculus
c	32	23	1.1	84767	9	AC004149	AC004149 Homo sapi
	33	23	1.1	145870	2	AC068675	AC068675 Homo sapi
c	34	23	1.1	170530	2	AL833805	AL833805 Mus muscu
c	35	23	1.1	177744	2	AC073954	AC073954 Homo sapi
	36	23	1.1	208955	2	AL808130	AL808130 Mus muscu
	37	22	1.0	1512	5	AF479823	AF479823 Pantodon
	38	22	1.0	43566	9	AF119709	AF119709 Homo sapi
	39	22	1.0	43922	9	HS73E16	Z95330 Human DNA s
	40	22	1.0	55237	2	AC119270	AC119270 Mus muscu
	41	22	1.0	163640	9	AC093328	AC093328 Homo sapi
	42	22	1.0	167794	2	AF215844	AF215844 Homo sapi
	43	22	1.0	185319	2	AC110364	AC110364 Rattus no
	44	22	1.0	237878	2	AC125202	AC125202 Mus muscu
	45	21	1.0	1198	5	AF251800	AF251800 Danio dan
	46	21	1.0	2492	5	DRY12007	Y12007 D.rerio vlg
	47	21	1.0	2865	5	AB005147	AB005147 Danio rer
	48	21	1.0	10010	1	U32709	U32709 Haemophilus
	49	21	1.0	70233	2	AC010831	AC010831 Homo sapi
c	50	21	1.0	84016	9	AL591375	AL591375 Human DNA
c	51	21	1.0	100288	2	AP001945	AP001945 Homo sapi
c	52	21	1.0	100853	9	AC094086	AC094086 Homo sapi
c	53	21	1.0	114724	2	AP004094	AP004094 Oryza sat
	54	21	1.0	116773	2	AC099186	AC099186 Rattus no
c	55	21	1.0	123593	9	AP001952	AP001952 Homo sapi
c	56	21	1.0	123661	9	AC096661	AC096661 Homo sapi
c	57	21	1.0	133695	2	AP003975	AP003975 Oryza sat
	58	21	1.0	138159	2	AP002742	AP002742 Homo sapi
	59	21	1.0	151019	9	AC026796	AC026796 Homo sapi
c	60	21	1.0	151663	2	AC102402	AC102402 Mus muscu
	61	21	1.0	157814	2	AC016078	AC016078 Homo sapi
	62	21	1.0	163793	2	AC109056	AC109056 Rattus no
c	63	21	1.0	167348	2	AP005104	AP005104 Oryza sat
c	64	21	1.0	172058	2	AC112815	AC112815 Rattus no
c	65	21	1.0	173316	2	AC110304	AC110304 Rattus no

	66	21	1.0	174627	9	AC022138	AC022138 Homo sapi
c	67	21	1.0	179144	9	AC026746	AC026746 Homo sapi
c	68	21	1.0	180303	2	AC018969	AC018969 Homo sapi
c	69	21	1.0	184826	2	AC122957	AC122957 Rattus no
c	70	21	1.0	185825	10	AL671868	AL671868 Mouse DNA
c	71	21	1.0	196605	2	AC114368	AC114368 Rattus no
c	72	21	1.0	196716	2	AC026750	AC026750 Homo sapi
	73	21	1.0	213351	2	AC026809	AC026809 Homo sapi
	74	20	0.9	309	8	SVU78623	U78623 Sesbania ve
	75	20	0.9	453	8	BSU78833	U78833 Bolusanthus
	76	20	0.9	630	8	CSU78838	U78838 Carmichaeli
c	77	20	0.9	714	11	PM12G5G	AL684440 Penicilli
	78	20	0.9	737	10	MUSPRHA	L25337 Mus musculu
c	79	20	0.9	1724	8	AF106843	AF106843 Peperomia
	80	20	0.9	2319	8	AY049285	AY049285 Arabidops
	81	20	0.9	2320	8	AY088132	AY088132 Arabidops
	82	20	0.9	2384	3	AB047381	AB047381 Hydra mag
	83	20	0.9	2502	5	AF046043	AF046043 Xenopus l
c	84	20	0.9	2624	9	HSEMP42G1	AF071540 Homo sapi
	85	20	0.9	9832	6	AX345559	AX345559 Sequence
	86	20	0.9	18946	3	CEC27D8	Z80214 Caenorhabdi
	87	20	0.9	25201	2	AC092204	AC092204 Homo sapi
	88	20	0.9	25347	3	CEC18D4	Z81474 Caenorhabdi
c	89	20	0.9	34841	3	CEC46F11	Z81449 Caenorhabdi
	90	20	0.9	38400	9	AC004058	AC004058 Homo sapi
c	91	20	0.9	56719	9	AC025466	AC025466 Homo sapi
c	92	20	0.9	56886	2	AC103759	AC103759 Homo sapi
c	93	20	0.9	59315	2	AC084169	AC084169 Homo sapi
c	94	20	0.9	61880	2	AC113129	AC113129 Mus muscu
	95	20	0.9	63910	10	AL646045	AL646045 Mouse DNA
c	96	20	0.9	64341	10	AL627323	AL627323 Mouse DNA
	97	20	0.9	64371	2	AC100463	AC100463 Mus muscu
	98	20	0.9	70540	9	AL136224	AL136224 Human DNA
	99	20	0.9	74712	2	AC121278	AC121278 Mus muscu
	100	20	0.9	89159	5	AL593845	AL593845 Zebrafish
	101	20	0.9	90370	2	AC112325	AC112325 Rattus no
c	102	20	0.9	92346	9	AC074289	AC074289 Homo sapi
	103	20	0.9	93398	2	AP001863	AP001863 Homo sapi
	104	20	0.9	94239	8	ATF14P22	AL137082 Arabidops
c	105	20	0.9	103517	8	AC073395	AC073395 Arabidops
c	106	20	0.9	106108	2	AC107129	AC107129 Rattus no
	107	20	0.9	107148	2	AC117003	AC117003 Rattus no
	108	20	0.9	107719	2	AC110847	AC110847 Rattus no
	109	20	0.9	108127	2	AC125477	AC125477 Medicago
c	110	20	0.9	108388	9	AP001821	AP001821 Homo sapi
	111	20	0.9	108451	2	AP002739	AP002739 Homo sapi
	112	20	0.9	109082	2	AC026133	AC026133 Homo sapi
	113	20	0.9	110282	9	AC096563	AC096563 Homo sapi
c	114	20	0.9	113028	9	HS68D15	AL049563 Human DNA
	115	20	0.9	118067	9	AC114942	AC114942 Homo sapi
	116	20	0.9	118777	10	AF163865	AF163865 Mus muscu
	117	20	0.9	119679	2	AC095884	AC095884 Rattus no
c	118	20	0.9	123905	9	AC116311	AC116311 Homo sapi
c	119	20	0.9	125607	2	AP003913	AP003913 Oryza sat
	120	20	0.9	127701	2	AC091302	AC091302 Oryza sat
c	121	20	0.9	129584	2	AC101911	AC101911 Mus muscu
c	122	20	0.9	130461	2	AP002741	AP002741 Homo sapi
c	123	20	0.9	132179	9	AL672142	AL672142 Human DNA
	124	20	0.9	133515	2	AL596091	AL596091 Homo sapi
	125	20	0.9	134922	9	AL589678	AL589678 Human DNA
c	126	20	0.9	136257	9	AC068797	AC068797 Homo sapi
c	127	20	0.9	139273	9	AL596275	AL596275 Human DNA
	128	20	0.9	143179	2	AC109346	AC109346 Homo sapi
	129	20	0.9	144303	2	AC108758	AC108758 Oryza sat
	130	20	0.9	146215	2	AC103617	AC103617 Mus muscu
	131	20	0.9	152883	8	AC084319	AC084319 Oryza sat
c	132	20	0.9	154286	9	AC016698	AC016698 Homo sapi
c	133	20	0.9	154535	9	AC114283	AC114283 Homo sapi
c	134	20	0.9	155539	9	AL138721	AL138721 Human DNA
c	135	20	0.9	155779	2	AC068260	AC068260 Homo sapi
c	136	20	0.9	156956	9	AC010080	AC010080 Homo sapi
	137	20	0.9	157165	2	AC104112	AC104112 Homo sapi
c	138	20	0.9	160324	9	AC079117	AC079117 Homo sapi

	139	20	0.9	161010	2	AC103631	AC103631 Mus muscu
	140	20	0.9	162897	2	AC128474	AC128474 Rattus no
c	141	20	0.9	162955	2	AC092877	AC092877 Homo sapi
	142	20	0.9	163699	2	AC112462	AC112462 Rattus no
c	143	20	0.9	163992	9	AC078889	AC078889 Homo sapi
	144	20	0.9	164229	9	AC006386	AC006386 Homo sapi
c	145	20	0.9	164808	9	HSDJ293L8	AL078594 Human DNA
	146	20	0.9	165474	2	AC025603	AC025603 Homo sapi
c	147	20	0.9	168081	2	AC083955	AC083955 Mus muscu
	148	20	0.9	169493	9	AL356605	AL356605 Human DNA
	149	20	0.9	169988	9	AC064846	AC064846 Homo sapi
c	150	20	0.9	172919	9	AL807742	AL807742 Human DNA
c	151	20	0.9	173377	2	AC114043	AC114043 Rattus no
	152	20	0.9	173852	2	AC095235	AC095235 Rattus no
	153	20	0.9	175601	2	AC103527	AC103527 Rattus no
c	154	20	0.9	175918	2	AC119617	AC119617 Rattus no
	155	20	0.9	177424	2	AC009832	AC009832 Homo sapi
	156	20	0.9	178660	9	AC009852	AC009852 Homo sapi
	157	20	0.9	180273	9	CNS01DU0	AL132994 Human chr
c	158	20	0.9	180797	9	AC116358	AC116358 Homo sapi
c	159	20	0.9	181716	2	AC113230	AC113230 Sus scrof
	160	20	0.9	182660	2	AC012416	AC012416 Homo sapi
c	161	20	0.9	182972	2	AC115454	AC115454 Rattus no
c	162	20	0.9	183228	2	AC102671	AC102671 Mus muscu
	163	20	0.9	183413	2	AC131200	AC131200 Rattus no
c	164	20	0.9	183719	2	AC068842	AC068842 Homo sapi
	165	20	0.9	183982	2	AC009825	AC009825 Homo sapi
c	166	20	0.9	184231	2	AC023651	AC023651 Homo sapi
c	167	20	0.9	184869	2	AC109695	AC109695 Rattus no
	168	20	0.9	185237	9	AC011124	AC011124 Homo sapi
	169	20	0.9	185286	10	AC006508	AC006508 Mus muscu
	170	20	0.9	185755	9	AC105460	AC105460 Homo sapi
c	171	20	0.9	185931	2	AL133554	AL133554 Homo sapi
c	172	20	0.9	186908	2	AC023770	AC023770 Homo sapi
	173	20	0.9	187999	2	AC026928	AC026928 Homo sapi
	174	20	0.9	188107	2	AC113229	AC113229 Sus scrof
c	175	20	0.9	188116	10	AL732392	AL732392 Mouse DNA
c	176	20	0.9	189590	2	AC034284	AC034284 Mus muscu
	177	20	0.9	190727	9	AC068724	AC068724 Homo sapi
	178	20	0.9	194466	9	AL591046	AL591046 Human DNA
	179	20	0.9	195916	2	AC097079	AC097079 Rattus no
c	180	20	0.9	199421	2	AC051636	AC051636 Homo sapi
c	181	20	0.9	199654	2	AC010898	AC010898 Homo sapi
c	182	20	0.9	200000	2	AC005289	AC005289 Homo sapi
	183	20	0.9	202429	2	AC130160	AC130160 Rattus no
c	184	20	0.9	202490	2	AC121884	AC121884 Mus muscu
	185	20	0.9	202849	2	AC130630	AC130630 Rattus no
	186	20	0.9	207951	2	AC012264	AC012264 Homo sapi
	187	20	0.9	208035	2	AC006520	AC006520 Mus muscu
	188	20	0.9	209392	2	AC113124	AC113124 Mus muscu
	189	20	0.9	209817	9	CNS018OT	AL109628 Human chr
c	190	20	0.9	212560	2	AC103688	AC103688 Homo sapi
	191	20	0.9	214441	2	AC123403	AC123403 Rattus no
	192	20	0.9	216800	10	AL589744	AL589744 Mouse DNA
c	193	20	0.9	219062	2	AC102840	AC102840 Mus muscu
c	194	20	0.9	236723	2	AC068771	AC068771 Homo sapi
	195	20	0.9	241664	2	AC124504	AC124504 Mus muscu
	196	20	0.9	250050	1	AL591976	AL591976 Listeria
c	197	20	0.9	251773	2	AC079129	AC079129 Mus muscu
	198	20	0.9	269616	2	AC124183	AC124183 Mus muscu
	199	20	0.9	305346	2	AC125327	AC125327 Mus muscu
	200	20	0.9	319404	2	AC112130	AC112130 Homo sapi
c	201	20	0.9	325956	2	AL807385	AL807385 Mus muscu
c	202	19	0.9	363	9	CJU04446	U04446 Callithrix
c	203	19	0.9	365	11	HS321WF1	Z51369 H.sapiens (
	204	19	0.9	489	8	AB076108S1	AB076108 Volvox ca
	205	19	0.9	489	8	AB076110S1	AB076110 Volvox ca
	206	19	0.9	599	3	CTI279788	AJ279788 Chrysolin
c	207	19	0.9	765	3	MSE6029	AJ006029 Manduca s
c	208	19	0.9	927	3	AF143849	AF143849 Adineta v
	209	19	0.9	983	8	AB076115S1	AB076115 Gonium mu
	210	19	0.9	1122	6	AX110914	AX110914 Sequence
	211	19	0.9	1128	8	AB013999	AB013999 Volvox ca

	212	19	0.9	1128	8	AB014006	AB014006 Pleodorin
	213	19	0.9	1128	8	AB014007	AB014007 Eudorina
	214	19	0.9	1128	8	AB014008	AB014008 Eudorina
	215	19	0.9	1128	8	AB014009	AB014009 Eudorina
	216	19	0.9	1128	8	AB014010	AB014010 Eudorina
	217	19	0.9	1128	8	AB014013	AB014013 Eudorina
	218	19	0.9	1128	8	AB014029	AB014029 Volvulina
	219	19	0.9	1128	8	AB014033	AB014033 Eudorina
	220	19	0.9	1128	8	AB044172	AB044172 Yamagishi
	221	19	0.9	1128	8	AB044533	AB044533 Lobomonas
	222	19	0.9	1128	8	AB047068	AB047068 Eudorina
	223	19	0.9	1128	8	AB047069	AB047069 Eudorina
	224	19	0.9	1128	8	AB047070	AB047070 Eudorina
	225	19	0.9	1128	8	AB047071	AB047071 Eudorina
	226	19	0.9	1128	8	AB047072	AB047072 Eudorina
	227	19	0.9	1128	8	AB047073	AB047073 Eudorina
	228	19	0.9	1128	8	AB076112	AB076112 Volvox gi
	229	19	0.9	1128	8	AB076113	AB076113 Volvox ob
	230	19	0.9	1128	8	AB076114	AB076114 Volvox af
	231	19	0.9	1128	8	AB076120	AB076120 Vitreochl
	232	19	0.9	1128	8	AB076121	AB076121 Vitreochl
	233	19	0.9	1128	8	AB076122	AB076122 Vitreochl
c	234	19	0.9	1132	10	BC023073	BC023073 Mus muscu
	235	19	0.9	1290	6	AX109755	AX109755 Sequence
	236	19	0.9	1291	6	AX109756	AX109756 Sequence
	237	19	0.9	1337	9	AK024583	AK024583 Homo sapi
c	238	19	0.9	1479	6	AX194078	AX194078 Sequence
	239	19	0.9	1586	5	AF520608	AF520608 Sparus au
	240	19	0.9	1710	10	BC009141	BC009141 Mus muscu
	241	19	0.9	1752	8	AF096248	AF096248 Lycopersi
	242	19	0.9	1898	9	AK001245	AK001245 Homo sapi
	243	19	0.9	2019	6	AX489264	AX489264 Sequence
	244	19	0.9	2028	9	AK093661	AK093661 Homo sapi
c	245	19	0.9	2123	9	AK027439	AK027439 Homo sapi
	246	19	0.9	2201	9	AF061337	AF061337 Homo sapi
c	247	19	0.9	2293	9	AK027471	AK027471 Homo sapi
c	248	19	0.9	2306	9	AK027654	AK027654 Homo sapi
	249	19	0.9	2319	9	HSAF000985	AF000985 Homo sapi
	250	19	0.9	2388	10	BC030895	BC030895 Mus muscu
	251	19	0.9	2487	9	AK027425	AK027425 Homo sapi
	252	19	0.9	2508	9	BC011819	BC011819 Homo sapi
	253	19	0.9	2556	9	AK091616	AK091616 Homo sapi
	254	19	0.9	2722	6	AX327460	AX327460 Sequence
	255	19	0.9	2838	10	AF159131	AF159131 Mus muscu
c	256	19	0.9	3041	9	AK027501	AK027501 Homo sapi
c	257	19	0.9	3093	6	AX026741	AX026741 Sequence
	258	19	0.9	3123	9	AK001652	AK001652 Homo sapi
	259	19	0.9	3158	9	HSU50553	U50553 Homo sapien
	260	19	0.9	3408	9	HSAF000983	AF000983 Homo sapi
c	261	19	0.9	3550	6	AX099520	AX099520 Sequence
c	262	19	0.9	3634	9	AK096337	AK096337 Homo sapi
c	263	19	0.9	3862	6	AX026746	AX026746 Sequence
c	264	19	0.9	3862	9	AF279656	AF279656 Homo sapi
c	265	19	0.9	4250	9	AB037789	AB037789 Homo sapi
	266	19	0.9	4308	3	AY070840	AY070840 Drosophil
	267	19	0.9	4416	9	HSAF000984	AF000984 Homo sapi
c	268	19	0.9	5197	8	LES6379	AJ006379 Lycopersi
	269	19	0.9	5322	9	HSAF000982	AF000982 Homo sapi
c	270	19	0.9	6060	9	AK027867	AK027867 Homo sapi
c	271	19	0.9	6073	2	AC012949	AC012949 Drosophil
c	272	19	0.9	7130	6	AX175460	AX175460 Sequence
c	273	19	0.9	7130	14	AF356698	AF356698 Porcine e
	274	19	0.9	7161	9	AB011168	AB011168 Homo sapi
	275	19	0.9	7174	9	HSM804580	AL833267 Homo sapi
	276	19	0.9	7238	3	CECLB2G	X56979 C. elegans
	277	19	0.9	9653	2	AC020842	AC020842 Mus muscu
	278	19	0.9	10029	1	AE010558	AE010558 Fusobacte
	279	19	0.9	10252	1	AE000075	AE000075 Rhizobium
	280	19	0.9	10477	1	AE007510	AE007510 Streptoco
	281	19	0.9	10478	1	AE008567	AE008567 Streptoco
c	282	19	0.9	10753	6	AX111980	AX111980 Sequence
c	283	19	0.9	10753	6	AX175459	AX175459 Sequence
c	284	19	0.9	12813	1	AE009776	AE009776 Pyrobacul

	285	19	0.9	13891	3	AC006665	AC006665 Caenorhab
	286	19	0.9	14087	1	AE006177	AE006177 Pasteurel
	287	19	0.9	14560	10	AB011527	AB011527 Rattus no
	288	19	0.9	18705	3	CEY26E6A	AL023830 Caenorhab
	289	19	0.9	18909	10	MMU223837	AJ223837 Mus muscu
	290	19	0.9	21230	9	HS381G6	AL031598 Human DNA
	291	19	0.9	23703	10	AF220365	AF220365 Mus muscu
c	292	19	0.9	27970	8	AP004518	AP004518 Lotus jap
c	293	19	0.9	29987	8	SC8156	Z49260 S.cerevisia
	294	19	0.9	31001	3	CEW07G4	Z78018 Caenorhabdi
c	295	19	0.9	32193	9	AC105348	AC105348 Homo sapi
c	296	19	0.9	33930	3	CEK04H4	Z27078 Caenorhabdi
c	297	19	0.9	35250	8	AY007366	AY007366 Lycopersi
	298	19	0.9	36948	2	AC014499	AC014499 Drosophil
	299	19	0.9	38968	9	AC005793	AC005793 Homo sapi
c	300	19	0.9	42074	9	AC004202	AC004202 Homo sapi
c	301	19	0.9	45977	2	AC020130	AC020130 Drosophil
c	302	19	0.9	46419	9	AC000058	AC000058 Homo sapi
c	303	19	0.9	48768	9	AL157933	AL157933 Human DNA
c	304	19	0.9	50000	6	AX393458	AX393458 Sequence
	305	19	0.9	51417	2	AC119930	AC119930 Mus muscu
	306	19	0.9	56835	2	AC025369	AC025369 Homo sapi
c	307	19	0.9	57339	2	AC044852	AC044852 Homo sapi
c	308	19	0.9	58484	9	AL137160	AL137160 Human DNA
c	309	19	0.9	58623	2	AC101064	AC101064 Mus muscu
	310	19	0.9	60310	9	AL391647	AL391647 Human DNA
	311	19	0.9	60514	2	AC014121	AC014121 Drosophil
	312	19	0.9	60555	2	AC116907	AC116907 Homo sapi
c	313	19	0.9	63650	2	AC084342	AC084342 Homo sapi
	314	19	0.9	64923	9	AC002545	AC002545 Homo sapi
	315	19	0.9	65386	2	AC002532	AC002532 Homo sapi
c	316	19	0.9	65972	2	AC124827	AC124827 Mus muscu
	317	19	0.9	66065	2	AC090169	AC090169 Homo sapi
c	318	19	0.9	66353	2	AC095348	AC095348 Homo sapi
c	319	19	0.9	67201	9	AL136169	AL136169 Human DNA
	320	19	0.9	67246	2	AC105035	AC105035 Homo sapi
c	321	19	0.9	68970	2	AC090378	AC090378 Homo sapi
c	322	19	0.9	71468	2	AC102077	AC102077 Mus muscu
c	323	19	0.9	71548	2	AC097573	AC097573 Rattus no
c	324	19	0.9	71585	2	AC119959	AC119959 Mus muscu
	325	19	0.9	73252	8	AC021199	AC021199 Arabidops
	326	19	0.9	73259	2	AC084006	AC084006 Homo sapi
	327	19	0.9	74686	2	AC104280	AC104280 Oryza sat
c	328	19	0.9	74989	2	AC023333	AC023333 Homo sapi
	329	19	0.9	75605	2	AC111486	AC111486 Rattus no
c	330	19	0.9	77075	2	AC130514	AC130514 Rattus no
c	331	19	0.9	79526	9	AC025244	AC025244 Homo sapi
c	332	19	0.9	83059	9	AP001435	AP001435 Homo sapi
	333	19	0.9	83550	9	AL390918	AL390918 Human DNA
c	334	19	0.9	83673	9	HSDA79C13	AL117325 Human DNA
	335	19	0.9	84277	2	AC105679	AC105679 Rattus no
c	336	19	0.9	86362	2	AC129445	AC129445 Rattus no
	337	19	0.9	89609	2	AC106134	AC106134 Rattus no
c	338	19	0.9	90171	9	AC095040	AC095040 Homo sapi
	339	19	0.9	90623	9	AC090993	AC090993 Homo sapi
	340	19	0.9	90636	9	AF230637	AF230637 Homo sapi
c	341	19	0.9	91045	8	AP003371	AP003371 Oryza sat
	342	19	0.9	91640	2	AC123295	AC123295 Rattus no
	343	19	0.9	95105	9	AC016561	AC016561 Homo sapi
	344	19	0.9	97687	9	AC008642	AC008642 Homo sapi
c	345	19	0.9	98398	9	AL731533	AL731533 Human DNA
c	346	19	0.9	100000	9	AP000008	AP000008 Homo sapi
c	347	19	0.9	100000	9	AP000149	AP000149 Homo sapi
c	348	19	0.9	100000	9	AP000518	AP000518 Homo sapi
c	349	19	0.9	100516	2	HSS171M_3	Continuation (4 of
	350	19	0.9	101539	8	ATT5P19	AL163972 Arabidops
	351	19	0.9	101624	2	AC102968	AC102968 Rattus no
c	352	19	0.9	102370	9	HSJ976O13	AL117354 Human DNA
c	353	19	0.9	103392	9	AC093793	AC093793 Homo sapi
	354	19	0.9	103473	2	AC126952	AC126952 Rattus no
	355	19	0.9	104311	2	AC105867	AC105867 Rattus no
	356	19	0.9	106552	2	AL805913	AL805913 Homo sapi
	357	19	0.9	106582	9	AL360215	AL360215 Human DNA

	358	19	0.9	108965	9	AF191069	AF191069 Homo sapi
c	359	19	0.9	109473	9	AL513363	AL513363 Human DNA
c	360	19	0.9	109637	9	AC084032	AC084032 Homo sapi
	361	19	0.9	109864	9	AF096876	AF096876 Homo sapi
c	362	19	0.9	109902	9	HSDJ46C2	AL132775 Human DNA
c	363	19	0.9	110000	2	AC117400_1	Continuation (2 of
	364	19	0.9	110000	2	AC129176_0	AC129176 Mus muscu
c	365	19	0.9	110000	2	AC129176_0	AC129176 Mus muscu
c	366	19	0.9	110000	2	AC129176_1	Continuation (2 of
	367	19	0.9	110000	2	AC129176_2	Continuation (3 of
c	368	19	0.9	110000	2	AC129176_2	Continuation (3 of
	369	19	0.9	111311	9	HS279I9	AL033517 Human DNA
	370	19	0.9	111837	9	AC069383	AC069383 Homo sapi
	371	19	0.9	112292	9	AC084297	AC084297 Homo sapi
	372	19	0.9	113797	9	AL590628	AL590628 Human DNA
c	373	19	0.9	114420	2	AC109763	AC109763 Rattus no
c	374	19	0.9	115016	2	AC096836	AC096836 Rattus no
c	375	19	0.9	116614	2	AL512349	AL512349 Homo sapi
	376	19	0.9	118130	2	AC127133	AC127133 Rattus no
	377	19	0.9	118189	2	AL513496	AL513496 Homo sapi
	378	19	0.9	118335	8	AC006193	AC006193 Arabidops
c	379	19	0.9	118572	9	AL133408	AL133408 Human DNA
	380	19	0.9	119004	9	AC010230	AC010230 Homo sapi
c	381	19	0.9	119395	9	AC091862	AC091862 Homo sapi
	382	19	0.9	119483	9	AC005588	AC005588 Homo sapi
c	383	19	0.9	121448	9	AC092699	AC092699 Homo sapi
c	384	19	0.9	122672	9	AP000704	AP000704 Homo sapi
c	385	19	0.9	123436	9	AC009872	AC009872 Homo sapi
c	386	19	0.9	124050	8	AF474373	AF474373 Hordeum v
	387	19	0.9	124591	2	AC115457	AC115457 Rattus no
c	388	19	0.9	125032	9	AC007099	AC007099 Homo sapi
c	389	19	0.9	125413	2	AC026607	AC026607 Homo sapi
	390	19	0.9	125585	9	HSDJ62D2	AL121955 Human DNA
	391	19	0.9	125990	9	AC108734	AC108734 Homo sapi
	392	19	0.9	126928	2	AC096673	AC096673 Trypanoso
	393	19	0.9	127178	9	AC005160	AC005160 Homo sapi
c	394	19	0.9	127892	8	AF411809	AF411809 Lycopersi
c	395	19	0.9	129090	2	AP003931	AP003931 Oryza sat
c	396	19	0.9	130615	9	AC010395	AC010395 Homo sapi
c	397	19	0.9	131142	9	AC073156	AC073156 Homo sapi
	398	19	0.9	131234	9	HS172K2	Z84814 Human DNA s
c	399	19	0.9	131823	2	AC010497	AC010497 Homo sapi
	400	19	0.9	132388	2	AC108190	AC108190 Felis cat
c	401	19	0.9	132505	2	AC015393	AC015393 Drosophil
c	402	19	0.9	133237	9	AC005550	AC005550 Homo sapi
c	403	19	0.9	133418	9	AL353709	AL353709 Human DNA
	404	19	0.9	134349	9	AC010296	AC010296 Homo sapi
c	405	19	0.9	135248	9	AC012610	AC012610 Homo sapi
	406	19	0.9	136225	9	AC004860	AC004860 Homo sapi
c	407	19	0.9	136682	2	AC105745	AC105745 Oryza sat
	408	19	0.9	137023	2	AC092736	AC092736 Trypanoso
	409	19	0.9	137307	2	AC087318	AC087318 Homo sapi
c	410	19	0.9	137610	9	AC013478	AC013478 Homo sapi
c	411	19	0.9	137849	2	AC105141	AC105141 Homo sapi
	412	19	0.9	138289	2	AP004873	AP004873 Oryza sat
	413	19	0.9	138594	9	AC107964	AC107964 Homo sapi
	414	19	0.9	138604	9	AC104826	AC104826 Homo sapi
	415	19	0.9	138905	2	AC131507	AC131507 Strongylo
	416	19	0.9	139350	2	AC097905	AC097905 Rattus no
	417	19	0.9	139376	9	AC020659	AC020659 Homo sapi
	418	19	0.9	139627	9	AL357117	AL357117 Human DNA
c	419	19	0.9	139644	2	AC123365	AC123365 Rattus no
	420	19	0.9	141040	8	AP004382	AP004382 Oryza sat
c	421	19	0.9	141159	2	AP000754	AP000754 Homo sapi
c	422	19	0.9	141356	2	AC128216	AC128216 Rattus no
	423	19	0.9	141748	8	AC009323	AC009323 Arabidops
	424	19	0.9	141925	2	AC098702	AC098702 Felis cat
c	425	19	0.9	142000	9	AP003110	AP003110 Homo sapi
	426	19	0.9	142362	2	AC036162	AC036162 Homo sapi
c	427	19	0.9	142600	9	AC005371	AC005371 Homo sapi
c	428	19	0.9	142667	8	AP003077	AP003077 Oryza sat
c	429	19	0.9	143116	2	AC023966	AC023966 Homo sapi
	430	19	0.9	143517	9	AC083783	AC083783 Homo sapi
c	431	19	0.9	143527	9	AC011353	AC011353 Homo sapi
c	432	19	0.9	143798	9	AL157409	AL157409 Human DNA
	433	19	0.9	144169	9	AL671859	AL671859 Human DNA
	434	19	0.9	144312	30	AL355822	Al355822 Human DNA
c	435	19	0.9	144622	2	AC121404	AC121404 Rattus no
	436	19	0.9	145638	9	HS357K22	AL022720 Human DNA
	437	19	0.9	145692	9	AL355303	AL355303 Human DNA
c	438	19	0.9	145734	2	AC110120	AC110120 Rattus no
	439	19	0.9	145739	8	OSJN00157	AL662957 Oryza sat
	440	19	0.9	146170	2	AC079224	AC079224 Homo sapi
	441	19	0.9	146190	2	AC073826	AC073826 Mus muscu
c	442	19	0.9	146377	9	AC083963	AC083963 Homo sapi
c	443	19	0.9	146936	2	AP005010	AP005010 Oryza sat
c	444	19	0.9	147078	2	AC080184	AC080184 Homo sapi
c	445	19	0.9	147122	9	AC090109	AC090109 Homo sapi
c	446	19	0.9	147358	9	AP002345	AP002345 Homo sapi
	447	19	0.9	147421	9	AC007857	AC007857 Homo sapi
c	448	19	0.9	147515	2	AC094016	AC094016 Pan trogl
	449	19	0.9	148280	9	AC004474	AC004474 Homo sapi
c	450	19	0.9	150022	8	AP003213	AP003213 Oryza sat
c	451	19	0.9	150305	2	AC114705	AC114705 Rattus no
	452	19	0.9	151514	2	AC120737	AC120737 Rattus no
	453	19	0.9	151740	2	AC096136	AC096136 Rattus no
c	454	19	0.9	152081	9	AC007381	AC007381 Homo sapi
	455	19	0.9	152334	9	AL356299	AL356299 Human DNA
c	456	19	0.9	152573	9	AL445202	AL445202 Human DNA
c	457	19	0.9	152666	2	AC128533	AC128533 Rattus no
c	458	19	0.9	152858	2	AC105881	AC105881 Rattus no
c	459	19	0.9	152975	2	AC129433	AC129433 Rattus no
	460	19	0.9	153198	9	AC026418	AC026418 Homo sapi
c	461	19	0.9	153468	9	AC113208	AC113208 Homo sapi
c	462	19	0.9	153879	2	AC094111	AC094111 Pan trogl
	463	19	0.9	153900	2	AC116408	AC116408 Mus muscu
c	464	19	0.9	154061	9	AC008524	AC008524 Homo sapi
	465	19	0.9	154284	2	AC116794	AC116794 Mus muscu
c	466	19	0.9	154926	9	AP005263	AP005263 Homo sapi
c	467	19	0.9	155518	9	AC015528	AC015528 Homo sapi
	468	19	0.9	155702	2	AC108196	AC108196 Felis cat
	469	19	0.9	155853	2	AC112965	AC112965 Mus muscu
	470	19	0.9	156325	9	HS134N8	AL031655 Human DNA
	471	19	0.9	156332	10	AC116581	AC116581 Mus muscu
c	472	19	0.9	156442	9	AL354915	AL354915 Human DNA
c	473	19	0.9	157183	9	AC011731	AC011731 Homo sapi
c	474	19	0.9	157235	3	AC009338	AC009338 Drosophil
	475	19	0.9	157466	2	AC105722	AC105722 Rattus no
c	476	19	0.9	157812	2	AC055805	AC055805 Homo sapi
c	477	19	0.9	158311	2	AC009589	AC009589 Homo sapi
	478	19	0.9	158648	2	AC021853	AC021853 Homo sapi
c	479	19	0.9	159063	2	AC111208	AC111208 Rattus no
	480	19	0.9	159178	9	AL669914	AL669914 Human DNA
	481	19	0.9	159390	9	AC092823	AC092823 Homo sapi
c	482	19	0.9	159451	9	AC012618	AC012618 Homo sapi
c	483	19	0.9	159624	9	AC069292	AC069292 Homo sapi
c	484	19	0.9	159986	9	AC025069	AC025069 Homo sapi
c	485	19	0.9	160623	9	AL670296	AL670296 Human DNA
c	486	19	0.9	160733	8	OSJN00182	AL662981 Oryza sat
	487	19	0.9	160736	2	AC027477	AC027477 Homo sapi
c	488	19	0.9	160869	9	AC021824	AC021824 Homo sapi
	489	19	0.9	161134	2	AC119709	AC119709 Rattus no
	490	19	0.9	161154	2	AC117340	AC117340 Rattus no
c	491	19	0.9	161493	2	AC125893	AC125893 Rattus no
	492	19	0.9	161667	9	AC005483	AC005483 Homo sapi
	493	19	0.9	161871	9	AP002806	AP002806 Homo sapi
c	494	19	0.9	162355	2	AC092957	AC092957 Homo sapi
c	495	19	0.9	162383	9	AB023054	AB023054 Homo sapi
c	496	19	0.9	162383	9	AB023055	AB023055 Homo sapi
c	497	19	0.9	162493	2	AC073972	AC073972 Homo sapi
c	498	19	0.9	162752	10	AC121961	AC121961 Mus muscu
c	499	19	0.9	162797	2	AC023219	AC023219 Homo sapi
	500	19	0.9	162839	9	AL137846	AL137846 Human DNA
	501	19	0.9	162996	9	AC006441	AC006441 Homo sapi
c	502	19	0.9	163067	2	AC022959	AC022959 Homo sapi
	503	19	0.9	163120	2	AC012480	AC012480 Homo sapi

c	504	19	0.9	163384	9	AC009180	AC009180 Homo sapi
	505	19	0.9	163842	2	AC096224	AC096224 Rattus no
	506	19	0.9	163881	9	AC007436	AC007436 Homo sapi
c	507	19	0.9	163978	2	AC126066	AC126066 Rattus no
c	508	19	0.9	164343	2	AC024593	AC024593 Homo sapi
c	509	19	0.9	164743	2	AP005636	AP005636 Oryza sat
c	510	19	0.9	164773	2	AC091543	AC091543 Felis cat
c	511	19	0.9	164810	2	AC109166	AC109166 Mus muscu
c	512	19	0.9	165024	2	AC093707	AC093707 Pan trogl
c	513	19	0.9	165093	9	AC069444	AC069444 Homo sapi
c	514	19	0.9	165385	2	AC103649	AC103649 Mus muscu
	515	19	0.9	165639	2	AC099623	AC099623 Mus muscu
c	516	19	0.9	166421	8	AP003106	AP003106 Oryza sat
c	517	19	0.9	166960	9	AC117513	AC117513 Homo sapi
c	518	19	0.9	167132	9	AC069154	AC069154 Homo sapi
c	519	19	0.9	167234	9	AC009729	AC009729 Homo sapi
	520	19	0.9	167275	9	AC008782	AC008782 Homo sapi
	521	19	0.9	167658	2	AL772326	AL772326 Mus muscu
	522	19	0.9	167705	2	AC074252	AC074252 Homo sapi
	523	19	0.9	167945	2	AC129291	AC129291 Mus muscu
c	524	19	0.9	168105	9	AC093815	AC093815 Homo sapi
	525	19	0.9	168413	2	AC009877	AC009877 Homo sapi
c	526	19	0.9	168437	2	AC124948	AC124948 Homo sapi
c	527	19	0.9	168505	9	AC104457	AC104457 Homo sapi
	528	19	0.9	168545	9	HSA287B20	AL121896 Human DNA
c	529	19	0.9	168954	2	AC099706	AC099706 Mus muscu
c	530	19	0.9	169192	2	AC114672	AC114672 Mus muscu
c	531	19	0.9	169237	2	AP004869	AP004869 Oryza sat
c	532	19	0.9	169354	2	AC099598	AC099598 Mus muscu
	533	19	0.9	169816	2	AC118600	AC118600 Mus muscu
	534	19	0.9	169841	9	AL607043	AL607043 Human DNA
c	535	19	0.9	170058	2	AC124905	AC124905 Equus cab
	536	19	0.9	170139	2	AC107335	AC107335 Rattus no
c	537	19	0.9	170393	9	AC026473	AC026473 Homo sapi
c	538	19	0.9	170458	2	AC034127	AC034127 Homo sapi
c	539	19	0.9	170795	2	AC021346	AC021346 Homo sapi
c	540	19	0.9	170891	2	AC002118	AC002118 Homo sapi
c	541	19	0.9	170962	2	AL389927	AL389927 Homo sapi
	542	19	0.9	170970	9	AL512510	AL512510 Human DNA
	543	19	0.9	171316	2	AL591712	AL591712 Mus muscu
	544	19	0.9	171400	2	AC129698	AC129698 Rattus no
c	545	19	0.9	171502	2	AC068638	AC068638 Homo sapi
c	546	19	0.9	171589	2	AC073567	AC073567 Homo sapi
	547	19	0.9	171613	2	AC109435	AC109435 Rattus no
	548	19	0.9	172266	3	AC092190	AC092190 Drosophil
	549	19	0.9	172360	2	AC024552	AC024552 Homo sapi
	550	19	0.9	172724	2	AC109968	AC109968 Rattus no
c	551	19	0.9	172931	9	AC006581	AC006581 Homo sapi
	552	19	0.9	172965	2	AC023367	AC023367 Homo sapi
	553	19	0.9	172993	9	CNS01RH9	AL161665 Human chr
c	554	19	0.9	173113	2	AC092529	AC092529 Papio cyn
c	555	19	0.9	173133	2	AC016166	AC016166 Homo sapi
c	556	19	0.9	173201	9	AC008169	AC008169 Homo sapi
	557	19	0.9	173289	2	AC096462	AC096462 Rattus no
	558	19	0.9	173348	2	AC115499	AC115499 Rattus no
c	559	19	0.9	173450	9	AC091769	AC091769 Homo sapi
c	560	19	0.9	173508	9	AC006385	AC006385 Homo sapi
	561	19	0.9	173597	2	AC025749	AC025749 Homo sapi
c	562	19	0.9	173635	2	AC123432	AC123432 Rattus no
c	563	19	0.9	173677	9	AC084768	AC084768 Homo sapi
c	564	19	0.9	174393	2	AC011720	AC011720 Homo sapi
c	565	19	0.9	174496	2	AC073076	AC073076 Homo sapi
	566	19	0.9	175131	2	AC114820	AC114820 Mus muscu
c	567	19	0.9	175210	9	CNS07EG5	AL590328 Human chr
c	568	19	0.9	175372	9	AC009659	AC009659 Homo sapi
c	569	19	0.9	175466	2	AC016419	AC016419 Homo sapi
c	570	19	0.9	175594	9	AL162503	AL162503 Human DNA
	571	19	0.9	175647	2	AC105414	AC105414 Homo sapi
	572	19	0.9	175878	2	AC130047	AC130047 Rattus no
	573	19	0.9	175951	9	AC011998	AC011998 Homo sapi
c	574	19	0.9	176104	9	AC007163	AC007163 Homo sapi
c	575	19	0.9	176493	9	AC016682	AC016682 Homo sapi
	576	19	0.9	177102	9	HSJ1069C8	AL078623 Human DNA

c	577	19	0.9	177835	3	AC091127	AC091127 Drosophil
	578	19	0.9	178257	2	AC011781	AC011781 Homo sapi
	579	19	0.9	178719	2	AC114135	AC114135 Rattus no
c	580	19	0.9	178965	2	AC096167	AC096167 Rattus no
	581	19	0.9	178986	2	AC046200	AC046200 Homo sapi
c	582	19	0.9	179387	2	AC113689	AC113689 Rattus no
c	583	19	0.9	179647	2	AC108124	AC108124 Homo sapi
	584	19	0.9	179856	3	AC010067	AC010067 Drosophil
c	585	19	0.9	180293	2	AC108143	AC108143 Homo sapi
c	586	19	0.9	180304	2	AC012541	AC012541 Homo sapi
c	587	19	0.9	180773	9	AC092969	AC092969 Homo sapi
	588	19	0.9	181245	9	AC093881	AC093881 Homo sapi
	589	19	0.9	181512	9	AC116017	AC116017 Homo sapi
c	590	19	0.9	181786	2	AL845506	AL845506 Mus muscu
c	591	19	0.9	181827	2	AC055854	AC055854 Homo sapi
c	592	19	0.9	183054	2	AC116074	AC116074 Rattus no
	593	19	0.9	183084	2	AC131222	AC131222 Rattus no
c	594	19	0.9	183244	2	AC027188	AC027188 Homo sapi
c	595	19	0.9	183451	9	AC005972	AC005972 Homo sapi
c	596	19	0.9	183547	2	AC009551	AC009551 Homo sapi
	597	19	0.9	183659	9	AC068472	AC068472 Homo sapi
	598	19	0.9	183698	10	AL606464	AL606464 Mouse DNA
	599	19	0.9	183838	9	AC104332	AC104332 Homo sapi
	600	19	0.9	184024	2	AC127124	AC127124 Rattus no
	601	19	0.9	184172	2	AC114244	AC114244 Rattus no
	602	19	0.9	184268	2	AC011401	AC011401 Homo sapi
c	603	19	0.9	184285	9	AC007365	AC007365 Homo sapi
	604	19	0.9	184391	9	AL592043	AL592043 Human DNA
	605	19	0.9	184427	14	EHVU20824	U20824 Equine herp
c	606	19	0.9	184579	2	AL824708	AL824708 Danio rer
c	607	19	0.9	184841	9	AC016902	AC016902 Homo sapi
c	608	19	0.9	184869	2	AC015622	AC015622 Homo sapi
c	609	19	0.9	185051	2	AC127098	AC127098 Rattus no
	610	19	0.9	185121	2	AC069105	AC069105 Homo sapi
c	611	19	0.9	185507	2	AC107562	AC107562 Rattus no
	612	19	0.9	185528	2	AC113643	AC113643 Rattus no
c	613	19	0.9	185569	2	AP003099	AP003099 Homo sapi
	614	19	0.9	185666	2	AC048389	AC048389 Homo sapi
	615	19	0.9	185709	2	AC094466	AC094466 Rattus no
c	616	19	0.9	185821	2	AL844866	AL844866 Mus muscu
c	617	19	0.9	186049	2	AC117695	AC117695 Mus muscu
	618	19	0.9	186709	2	AC113713	AC113713 Rattus no
	619	19	0.9	187332	2	AC101292	AC101292 Mus muscu
	620	19	0.9	187359	2	AC068886	AC068886 Homo sapi
	621	19	0.9	187509	2	AC114370	AC114370 Rattus no
	622	19	0.9	187680	2	AC126244	AC126244 Mus muscu
	623	19	0.9	187908	9	HS495O10	AL031121 Human DNA
	624	19	0.9	187942	9	AC104639	AC104639 Homo sapi
	625	19	0.9	187976	2	AC127750	AC127750 Rattus no
	626	19	0.9	188018	2	AC021525	AC021525 Homo sapi
	627	19	0.9	188207	2	AC010233	AC010233 Homo sapi
c	628	19	0.9	188425	2	AC026554	AC026554 Homo sapi
	629	19	0.9	188697	2	AC125577	AC125577 Rattus no
c	630	19	0.9	188781	10	AL713870	AL713870 Mouse DNA
	631	19	0.9	189056	2	AC024519	AC024519 Homo sapi
c	632	19	0.9	189092	2	AC011759	AC011759 Drosophil
c	633	19	0.9	189355	2	AP000921	AP000921 Homo sapi
	634	19	0.9	189572	9	AC091199	AC091199 Homo sapi
c	635	19	0.9	189829	2	AC113069	AC113069 Mus muscu
c	636	19	0.9	190338	2	AC130145	AC130145 Rattus no
	637	19	0.9	190626	2	AC062036	AC062036 Homo sapi
c	638	19	0.9	191058	2	AC105910	AC105910 Homo sapi
	639	19	0.9	191704	2	AC053529	AC053529 Homo sapi
c	640	19	0.9	191728	2	AL512269	AL512269 Homo sapi
c	641	19	0.9	191923	2	AP001846	AP001846 Homo sapi
c	642	19	0.9	192399	3	AC010032	AC010032 Drosophil
c	643	19	0.9	192748	2	AC011089	AC011089 Homo sapi
c	644	19	0.9	192869	2	AC113326	AC113326 Mus muscu
	645	19	0.9	192967	2	AC102767	AC102767 Mus muscu
c	646	19	0.9	193302	2	AC021992	AC021992 Homo sapi
	647	19	0.9	193453	2	AC126553	AC126553 Mus muscu
	648	19	0.9	193761	2	AC022273	AC022273 Homo sapi
	649	19	0.9	193774	9	AL445183	AL445183 Human DNA

	650	19	0.9	193872	9	AC087369	AC087369 Homo sapi
c	651	19	0.9	193942	2	AC112925	AC112925 Mus muscu
c	652	19	0.9	195002	2	AC123812	AC123812 Mus muscu
	653	19	0.9	195296	2	AC122234	AC122234 Mus muscu
	654	19	0.9	195296	2	AC026606	AC026606 Homo sapi
c	655	19	0.9	195310	2	AC024189	AC024189 Homo sapi
	656	19	0.9	195680	9	AC120498	AC120498 Homo sapi
c	657	19	0.9	196363	9	AC096759	AC096759 Homo sapi
	658	19	0.9	196472	2	AC011862	AC011862 Homo sapi
	659	19	0.9	196636	2	AC120676	AC120676 Rattus no
c	660	19	0.9	197351	2	AC127204	AC127204 Rattus no
c	661	19	0.9	197403	2	AC123823	AC123823 Mus muscu
	662	19	0.9	198259	2	AC117648	AC117648 Mus muscu
	663	19	0.9	198582	9	AC005291	AC005291 Homo sapi
	664	19	0.9	198958	9	AC020688	AC020688 Homo sapi
c	665	19	0.9	199016	2	AC021387	AC021387 Homo sapi
c	666	19	0.9	200282	9	AP003100	AP003100 Homo sapi
c	667	19	0.9	200339	10	AC122833	AC122833 Mus muscu
c	668	19	0.9	200703	2	AC123482	AC123482 Rattus no
	669	19	0.9	200843	2	AC111119	AC111119 Mus muscu
c	670	19	0.9	200853	9	CNS01DSQ	AL121839 Human chr
	671	19	0.9	201136	9	AP000959	AP000959 Homo sapi
c	672	19	0.9	201312	2	AC067900	AC067900 Homo sapi
	673	19	0.9	201676	2	AC105367	AC105367 Sus scrof
c	674	19	0.9	202565	9	AL354696	AL354696 Human DNA
	675	19	0.9	205339	10	AC091484	AC091484 Mus muscu
c	676	19	0.9	205993	2	AC011952	AC011952 Homo sapi
c	677	19	0.9	207018	2	AC116539	AC116539 Drosophil
c	678	19	0.9	207162	2	AL831752	AL831752 Mus muscu
	679	19	0.9	207268	2	AL671854	AL671854 Mus muscu
c	680	19	0.9	207319	2	AC073854	AC073854 Homo sapi
c	681	19	0.9	207534	2	AC024698	AC024698 Homo sapi
	682	19	0.9	207791	9	AC007790	AC007790 Homo sapi
c	683	19	0.9	208265	10	AL606907	AL606907 Mouse DNA
c	684	19	0.9	208398	2	AC123245	AC123245 Rattus no
	685	19	0.9	208695	10	AL671478	AL671478 Mouse DNA
	686	19	0.9	208936	2	AC113181	AC113181 Mus muscu
c	687	19	0.9	209910	2	AC105301	AC105301 Mus muscu
c	688	19	0.9	210813	10	AL627070	AL627070 Mouse DNA
c	689	19	0.9	210933	2	AC027394	AC027394 Homo sapi
c	690	19	0.9	211357	2	AC021136	AC021136 Homo sapi
	691	19	0.9	211433	2	AL805896	AL805896 Mus muscu
c	692	19	0.9	211844	2	AC048370	AC048370 Homo sapi
c	693	19	0.9	212529	2	AC115948	AC115948 Mus muscu
	694	19	0.9	212560	2	AC103688	AC103688 Homo sapi
c	695	19	0.9	213477	2	AC098347	AC098347 Rattus no
	696	19	0.9	213943	2	AC015465	AC015465 Homo sapi
c	697	19	0.9	214572	10	AL450341	AL450341 Mouse DNA
c	698	19	0.9	214596	2	AC122233	AC122233 Mus muscu
	699	19	0.9	216457	9	CNS01RG1	AL157689 Human chr
c	700	19	0.9	217586	2	AC016076	AC016076 Homo sapi
c	701	19	0.9	218565	3	AC009257	AC009257 Drosophil
	702	19	0.9	218956	9	CNS00008	AL049831 Human chr
	703	19	0.9	219145	2	AC102087	AC102087 Mus muscu
c	704	19	0.9	219699	2	AC105206	AC105206 Homo sapi
c	705	19	0.9	219825	10	AC098886	AC098886 Mus muscu
	706	19	0.9	219973	10	AL670227	AL670227 Mouse DNA
c	707	19	0.9	220407	2	AC112792	AC112792 Mus muscu
c	708	19	0.9	220450	2	HS53I10	AL133493 Homo sapi
c	709	19	0.9	220568	9	AC108048	AC108048 Homo sapi
c	710	19	0.9	221399	10	AL604045	AL604045 Mouse DNA
	711	19	0.9	224360	2	AC090670	AC090670 Homo sapi
c	712	19	0.9	224613	2	AC126555	AC126555 Mus muscu
c	713	19	0.9	226905	3	AE003773	AE003773 Drosophil
c	714	19	0.9	229277	2	AC087140	AC087140 Mus muscu
c	715	19	0.9	230516	9	HSU52111	U52111 Homo sapien
	716	19	0.9	232807	2	SPNEU1901	AL449923 Streptoco
c	717	19	0.9	234053	9	AC002429	AC002429 Human BAC
	718	19	0.9	235183	2	AC079427	AC079427 Mus muscu
	719	19	0.9	235733	2	AC105297	AC105297 Mus muscu
	720	19	0.9	236809	2	AL356746	AL356746 Homo sapi
c	721	19	0.9	237931	9	AC022098	AC022098 Homo sapi
	722	19	0.9	238413	2	AC124513	AC124513 Mus muscu
c	723	19	0.9	238845	2	AC103330	AC103330 Rattus no
c	724	19	0.9	247876	2	AL844873	AL844873 Mus muscu
c	725	19	0.9	248381	2	AC009937	AC009937 Homo sapi
	726	19	0.9	249616	10	AE014183	AE014183 Mus muscu
c	727	19	0.9	259474	9	HUAC004605	AC004605 Homo sapi
	728	19	0.9	263692	2	AC113251	AC113251 Rattus no
	729	19	0.9	265786	9	AE006466	AE006466 Homo sapi
c	730	19	0.9	269616	2	AC124183	AC124183 Mus muscu
	731	19	0.9	270100	2	AC126428	AC126428 Mus muscu
c	732	19	0.9	274388	3	AE003574	AE003574 Drosophil
	733	19	0.9	287178	2	AC125377	AC125377 Mus muscu
	734	19	0.9	291050	1	AP000982	AP000982 Sulfolobu
c	735	19	0.9	295225	3	AE003461	AE003461 Drosophil
c	736	19	0.9	301711	3	AE003470	AE003470 Drosophil
	737	19	0.9	302300	1	AP003186	AP003186 Clostridi
c	738	19	0.9	309883	3	AE003479	AE003479 Drosophil
	739	19	0.9	317463	2	AC091254	AC091254 Mus muscu
	740	19	0.9	320040	6	A79350	A79350 Sequence 1
	741	19	0.9	320040	6	A93002	A93002 Sequence 1
c	742	19	0.9	320436	2	AL772140	AL772140 Homo sapi
	743	19	0.9	325808	2	AC121787	AC121787 Mus muscu
	744	19	0.9	340000	9	AP001684	AP001684 Homo sapi
	745	19	0.9	340000	9	AP001707	AP001707 Homo sapi
c	746	19	0.9	340000	9	AP001727	AP001727 Homo sapi
c	747	19	0.9	340000	9	HS21C102	AL163302 Homo sapi
	748	19	0.9	349980	6	AX344565	AX344565 Sequence
	749	19	0.9	349980	6	AX344566	AX344566 Sequence
	750	18	0.8	107	9	HS50D11R	Z61345 H.sapiens C
c	751	18	0.8	163	11	HUMSWX1698	L42703 Human chrom
c	752	18	0.8	174	6	AX088553	AX088553 Sequence
	753	18	0.8	250	14	AF106188	AF106188 HIV-1 str
	754	18	0.8	256	9	HS50B8R	Z65642 H.sapiens C
	755	18	0.8	339	6	AR203677	AR203677 Sequence
	756	18	0.8	340	3	AB047387	AB047387 Dugesia d
	757	18	0.8	390	14	AB081225	AB081225 Human imm
	758	18	0.8	390	14	AF015397	AF015397 HIV-1 str
	759	18	0.8	390	14	AF015398	AF015398 HIV-1 str
	760	18	0.8	390	14	AF015399	AF015399 HIV-1 str
	761	18	0.8	390	14	AF062042	AF062042 HIV-1 iso
	762	18	0.8	390	14	AF062043	AF062043 HIV-1 iso
	763	18	0.8	390	14	AF062044	AF062044 HIV-1 iso
	764	18	0.8	392	11	G21206	G21206 human STS W
	765	18	0.8	396	14	HIV1SIWMN	Z79544 Human immun
	766	18	0.8	421	6	AX302796	AX302796 Sequence
c	767	18	0.8	421	11	G07123	G07123 human STS W
	768	18	0.8	438	14	AF481781	AF481781 HIV-1 iso
	769	18	0.8	452	6	AX110906	AX110906 Sequence
	770	18	0.8	532	6	AR138292	AR138292 Sequence
	771	18	0.8	579	11	G66094	G66094 sY1028 Misc
c	772	18	0.8	599	6	AX331374	AX331374 Sequence
	773	18	0.8	710	9	HSA336229	AJ336229 Homo sapi
	774	18	0.8	717	11	PM4H3G	AL685947 Penicilli
	775	18	0.8	721	14	AF051470	AF051470 HIV-1 iso
	776	18	0.8	732	14	AF196674	AF196674 HIV-1 iso
	777	18	0.8	836	9	HUMIGHAJ	M12859 Human Ig ab
	778	18	0.8	882	11	CNS06EXZ	AL395709 T7 end of
c	779	18	0.8	895	3	AF027724	AF027724 Eimeria m
c	780	18	0.8	909	6	AX463545	AX463545 Sequence
c	781	18	0.8	912	8	AF404858	AF404858 Arabidops
c	782	18	0.8	957	6	AX241558	AX241558 Sequence
c	783	18	0.8	1004	11	CNS06F7M	AL396056 T7 end of
	784	18	0.8	1018	6	AX109947	AX109947 Sequence
c	785	18	0.8	1028	11	CNS06IMR	AL400489 T7 end of
	786	18	0.8	1105	3	AY119531	AY119531 Drosophil
	787	18	0.8	1125	5	AF110995	AF110995 Gallus ga
	788	18	0.8	1128	8	AB014000	AB014000 Volvox di
	789	18	0.8	1128	8	AB014019	AB014019 Gonium qu
	790	18	0.8	1128	8	AB014020	AB014020 Gonium mu
c	791	18	0.8	1132	3	AF308585	AF308585 Tritricho
c	792	18	0.8	1150	8	AY088401	AY088401 Arabidops
c	793	18	0.8	1195	3	AF529422	AF529422 Bombyx mo
	794	18	0.8	1222	8	CCY419134	AJ419134 Chorizand
	795	18	0.8	1263	10	AF238223	AF238223 Rattus no

	796	18	0.8	1277	6	AX109728	AX109728 Sequence
	797	18	0.8	1331	8	AY056519	AY056519 Encephala
	798	18	0.8	1337	8	AF209597	AF209597 Heteropyx
	799	18	0.8	1346	8	MSA248322	AJ248322 Medicago
	800	18	0.8	1348	8	PLI417571	AJ417571 Paradisea
	801	18	0.8	1352	8	SSP419158	AJ419158 Streptoch
	802	18	0.8	1369	8	FPR235477	AJ235477 Fuchsia p
	803	18	0.8	1374	8	AF209564	AF209564 Clarkia x
	804	18	0.8	1380	8	ORU93825	U93825 Ophioglossu
c	805	18	0.8	1410	1	SAU309184	AJ309184 Staphyloc
	806	18	0.8	1418	8	AF209621	AF209621 Lythrum s
	807	18	0.8	1423	10	AF250142	AF250142 Rattus no
	808	18	0.8	1426	8	AF209582	AF209582 Epilobium
c	809	18	0.8	1428	5	AF484684	AF484684 Anas plat
c	810	18	0.8	1428	10	AF317514	AF317514 Cavia por
c	811	18	0.8	1428	10	MMU10374	U10374 Mus musculu
	812	18	0.8	1429	8	AF209637	AF209637 Neurada p
	813	18	0.8	1432	8	AF197601	AF197601 Triglochi
	814	18	0.8	1433	8	ATH010458	AJ010458 Arabidops
	815	18	0.8	1438	8	PPR235574	AJ235574 Punica pr
	816	18	0.8	1439	8	AF209633	AF209633 Moringa o
	817	18	0.8	1443	8	AF209686	AF209686 Terminali
	818	18	0.8	1443	8	ALI235394	AJ235394 Anthericu
	819	18	0.8	1445	8	PPA419149	AJ419149 Pharus pa
	820	18	0.8	1453	8	EMO419136	AJ419136 Ecdeiocol
	821	18	0.8	1453	8	NPAJ3069	AJ233069 Neurada p
	822	18	0.8	1453	8	QIN235576	AJ235576 Quisquali
	823	18	0.8	1456	10	AF020039	AF020039 Mus muscu
	824	18	0.8	1464	8	AF197610	AF197610 Orontium
	825	18	0.8	1467	8	AAAJ3072	AJ233072 Abroma an
	826	18	0.8	1467	8	MNE235535	AJ235535 Metroside
	827	18	0.8	1471	8	AF197600	AF197600 Potamoget
	828	18	0.8	1474	8	AF209677	AF209677 Smilax gl
	829	18	0.8	1474	8	SWA235606	AJ235606 Spathiphy
	830	18	0.8	1475	8	JGA419143	AJ419143 Joinville
	831	18	0.8	1480	8	CEN419133	AJ419133 Chorizand
	832	18	0.8	1480	8	HNE419132	AJ419132 Hypolytru
	833	18	0.8	1491	8	AF168906	AF168906 Cyperus a
	834	18	0.8	1491	8	AF168915	AF168915 Gymnostac
	835	18	0.8	1495	14	HIVG109	L11765 Human immun
	836	18	0.8	1497	8	AAF417568	AJ417568 Agapanthu
	837	18	0.8	1500	8	AF168887	AF168887 Aponogeto
	838	18	0.8	1500	8	AF168894	AF168894 Chlorophy
	839	18	0.8	1500	8	AF168896	AF168896 Colocasia
	840	18	0.8	1500	8	AF168898	AF168898 Costus ba
	841	18	0.8	1500	8	AF168909	AF168909 Dimerocos
	842	18	0.8	1500	8	AF168930	AF168930 Monocostu
	843	18	0.8	1500	8	AF168949	AF168949 Tapeinoch
	844	18	0.8	1500	8	AF168950	AF168950 Tradescan
	845	18	0.8	1536	6	AR138293	AR138293 Sequence
	846	18	0.8	1536	9	HSTIF	X79538 H.sapiens n
	847	18	0.8	1553	9	HUMFGFRB	L78720 Human fibro
	848	18	0.8	1581	6	AX489060	AX489060 Sequence
c	849	18	0.8	1608	9	AF285158	AF285158 Homo sapi
	850	18	0.8	1624	14	AB074067	AB074067 Human imm
	851	18	0.8	1628	9	BC003662	BC003662 Homo sapi
	852	18	0.8	1635	14	AB074069	AB074069 Human imm
	853	18	0.8	1640	9	BC004386	BC004386 Homo sapi
	854	18	0.8	1640	9	BC011151	BC011151 Homo sapi
	855	18	0.8	1645	8	AY117298	AY117298 Arabidops
	856	18	0.8	1648	14	AB074068	AB074068 Human imm
	857	18	0.8	1674	10	MMU46690	U46690 Mus musculu
	858	18	0.8	1682	6	AR138294	AR138294 Sequence
	859	18	0.8	1682	9	HUMORFJA	D21853 Human mRNA
	860	18	0.8	1748	3	DDHEL2A	X81823 D.discoideu
	861	18	0.8	1779	5	GAG223618	AJ223618 Gallus ga
c	862	18	0.8	1780	10	MMU09138	U09138 Mus musculu
c	863	18	0.8	1782	10	BC021798	BC021798 Mus muscu
c	864	18	0.8	1800	10	U01664	U01664 Mus musculu
	865	18	0.8	1849	8	AY080680	AY080680 Arabidops
	866	18	0.8	1876	6	AR159782	AR159782 Sequence
	867	18	0.8	1876	6	BD010462	BD010462 SMAD7 and
	868	18	0.8	1914	6	AR208574	AR208574 Sequence

	869	18	0.8	1972	5	AF479820	AF479820 Cyprinus
	870	18	0.8	2003	10	RNSMAD4	AF156730 Rattus no
c	871	18	0.8	2005	6	AR030509	AR030509 Sequence
c	872	18	0.8	2085	1	AY099293	AY099293 Bartonell
	873	18	0.8	2105	1	STU94901	U94901 Salmonella
c	874	18	0.8	2130	10	MMU01841	U01841 Mus musculu
c	875	18	0.8	2141	8	AF105110	AF105110 Nymphaea
c	876	18	0.8	2196	6	AX088888	AX088888 Sequence
	877	18	0.8	2219	9	AB070050	AB070050 Macaca fa
c	878	18	0.8	2261	6	AX088887	AX088887 Sequence
	879	18	0.8	2317	3	AF004916	AF004916 Anopheles
	880	18	0.8	2428	3	AGA010195	AJ010195 Anopheles
	881	18	0.8	2480	5	AF134850	AF134850 Danio rer
c	882	18	0.8	2500	9	HUMCGMP	M84741 Human cGMP-
	883	18	0.8	2501	9	HSSMAD7S04	AF026559 Homo sapi
c	884	18	0.8	2548	8	AY102111	AY102111 Arabidops
	885	18	0.8	2632	14	AB078707	AB078707 Human imm
c	886	18	0.8	2715	6	AX399012	AX399012 Sequence
	887	18	0.8	2715	8	PTOCPRGOL	L42851 Prototheca
	888	18	0.8	2812	5	XLU92030	U92030 Xenopus lae
c	889	18	0.8	2857	9	S42457	S42457 CNCG=rod ph
c	890	18	0.8	2871	9	HUMB2M02	M17987 Homo sapien
	891	18	0.8	2902	10	AF214732	AF214732 Mus muscu
	892	18	0.8	2910	3	AY075460	AY075460 Drosophil
	893	18	0.8	2955	10	AF464181	AF464181 Mus muscu
c	894	18	0.8	3029	9	HSU88879	U88879 Homo sapien
	895	18	0.8	3103	6	AR103433	AR103433 Sequence
	896	18	0.8	3103	6	AR201459	AR201459 Sequence
	897	18	0.8	3111	6	AR121107	AR121107 Sequence
	898	18	0.8	3111	6	AX334395	AX334395 Sequence
	899	18	0.8	3111	9	AF010193	AF010193 Homo sapi
	900	18	0.8	3157	8	S51252	S51252 Neurospora
	901	18	0.8	3165	3	AY089433	AY089433 Drosophil
	902	18	0.8	3222	8	ATU83178	U83178 Arabidopsis
c	903	18	0.8	3321	6	BD004082	BD004082 Apoptosis
c	904	18	0.8	3513	6	AX277590	AX277590 Sequence
c	905	18	0.8	3513	9	HSM800944	AL117435 Homo sapi
c	906	18	0.8	3522	5	XELSTAGE	L04278 Xenopus lae
	907	18	0.8	3681	10	MMSMAD7	AJ000550 Mus muscu
c	908	18	0.8	4046	9	HSM800935	AL117393 Homo sapi
	909	18	0.8	4311	10	AF159626	AF159626 Rattus no
c	910	18	0.8	4477	9	AK024475	AK024475 Homo sapi
	911	18	0.8	4528	10	AY013805	AY013805 Mus muscu
	912	18	0.8	5001	8	CHPW5KB	AJ222802 Protothec
	913	18	0.8	5143	5	AF067418	AF067418 Xenopus l
	914	18	0.8	5714	2	AC015032	AC015032 Drosophil
	915	18	0.8	6608	1	AB070346	AB070346 Tetrageno
c	916	18	0.8	6763	6	BD004101	BD004101 Apoptosis
c	917	18	0.8	6763	10	AF132726	AF132726 Mus muscu
c	918	18	0.8	6927	9	AB011103	AB011103 Homo sapi
	919	18	0.8	7098	2	AC023900	AC023900 Mus muscu
c	920	18	0.8	7292	8	CAEFGTF	Z32687 Candida alb
c	921	18	0.8	7403	1	STU69493	U69493 Salmonella
	922	18	0.8	7582	14	AF492770	AF492770 Sindbis v
	923	18	0.8	8414	9	AY038064	AY038064 Homo sapi
	924	18	0.8	8491	6	AX449276	AX449276 Sequence
	925	18	0.8	8491	6	AX449278	AX449278 Sequence
	926	18	0.8	8491	9	HUMHH1A	M77235 Human cardi
	927	18	0.8	8922	1	AE010235	AE010235 Pyrococcu
c	928	18	0.8	9099	8	AF315315	AF315315 Nectria h
	929	18	0.8	9190	14	HIVOYI	M26727 Human immun
	930	18	0.8	9771	1	U67492	U67492 Methanococc
c	931	18	0.8	10029	1	AE006797	AE006797 Sulfolobu
c	932	18	0.8	10039	1	AE013108	AE013108 Thermoana
	933	18	0.8	10082	4	AY071822	AY071822 Bos tauru
	934	18	0.8	10137	1	AE013296	AE013296 Methanosa
	935	18	0.8	10183	1	AE007790	AE007790 Clostridi
c	936	18	0.8	10547	1	AE010506	AE010506 Fusobacte
	937	18	0.8	10549	1	U67511	U67511 Methanococc
	938	18	0.8	10591	1	AE013379	AE013379 Methanosa
	939	18	0.8	10691	1	AE010530	AE010530 Fusobacte
	940	18	0.8	11014	1	AE010501	AE010501 Fusobacte
c	941	18	0.8	11026	8	AB020753	AB020753 Arabidops

	942	18	0.8	11956	1	AE006094	AE006094 Pasteurel
c	943	18	0.8	12436	1	AE011894	AE011894 Xanthomon
c	944	18	0.8	12637	1	AF411935	AF411935 Staphyloc
c	945	18	0.8	15549	1	AE001482	AE001482 Helicobac
c	946	18	0.8	15556	3	AE002909	AE002909 Drosophil
	947	18	0.8	15871	8	AB010696	AB010696 Arabidops
c	948	18	0.8	16325	2	AC015144	AC015144 Drosophil
	949	18	0.8	16662	3	AC114262	AC114262 Dictyoste
	950	18	0.8	16671	3	AF038604	AF038604 Caenorhab
	951	18	0.8	17747	2	AC017186	AC017186 Drosophil
c	952	18	0.8	19024	3	U00038	U00038 Caenorhabdi
	953	18	0.8	20951	1	AE008715	AE008715 Salmonell
c	954	18	0.8	21649	8	SPCP31B10	AL121859 S.pombe c
	955	18	0.8	22085	8	PWI245645	AJ245645 Protothec
	956	18	0.8	24440	5	DRE311625	AJ311625 Danio rer
	957	18	0.8	24470	2	AC095802	AC095802 Rattus no
	958	18	0.8	25000	5	AF461759	AF461759 Danio rer
c	959	18	0.8	25842	3	CEF57F5	Z75953 Caenorhabdi
	960	18	0.8	26600	3	CEF37B1	Z80217 Caenorhabdi
	961	18	0.8	27166	8	SC9531X	Z49919 S.cerevisia
	962	18	0.8	27178	3	U88175	U88175 Caenorhabdi
c	963	18	0.8	28163	3	HTU252741	AJ252741 Haliotis
c	964	18	0.8	31001	3	CEW07G4	Z78018 Caenorhabdi
	965	18	0.8	31196	2	AC020371	AC020371 Drosophil
	966	18	0.8	32063	3	CECC4	Z81490 Caenorhabdi
c	967	18	0.8	34324	2	AC102844	AC102844 Mus muscu
c	968	18	0.8	35542	9	AL603825	AL603825 Human DNA
c	969	18	0.8	36335	9	AC011349	AC011349 Homo sapi
	970	18	0.8	37535	9	AC113341	AC113341 Homo sapi
c	971	18	0.8	37860	3	AC022284	AC022284 Leishmani
	972	18	0.8	38478	2	AC019893	AC019893 Drosophil
	973	18	0.8	38779	8	SCU31900	U31900 Saccharomyc
	974	18	0.8	38812	3	AC092775	AC092775 Leishmani
	975	18	0.8	38825	9	AC090098	AC090098 Homo sapi
	976	18	0.8	39225	9	AC072053	AC072053 Homo sapi
c	977	18	0.8	39989	3	CEF14F8	Z92782 Caenorhabdi
c	978	18	0.8	40014	4	AC091538	AC091538 Sus scrof
	979	18	0.8	40020	3	AC068783	AC068783 Leishmani
c	980	18	0.8	40080	2	AC011563	AC011563 Homo sapi
c	981	18	0.8	40356	9	AC118063	AC118063 Homo sapi
c	982	18	0.8	42021	3	CEC05D12	Z50872 Caenorhabdi
c	983	18	0.8	42056	3	CEC53D6	Z70270 Caenorhabdi
c	984	18	0.8	42569	9	HS305C8	AL031705 Human DNA
c	985	18	0.8	43657	9	AL499626	AL499626 Human DNA
c	986	18	0.8	44419	9	AC006506	AC006506 Homo sapi
c	987	18	0.8	45819	9	AC091650	AC091650 Homo sapi
c	988	18	0.8	46185	9	AL157782	AL157782 Human DNA
c	989	18	0.8	46197	2	AC104764	AC104764 Homo sapi
c	990	18	0.8	50175	9	AL358393	AL358393 Human DNA
	991	18	0.8	52095	10	AL772180	AL772180 Mouse DNA
	992	18	0.8	52514	2	AC116551	AC116551 Dictyoste
	993	18	0.8	52780	9	AC026762	AC026762 Homo sapi
c	994	18	0.8	53135	9	AC116621	AC116621 Homo sapi
c	995	18	0.8	53867	10	AL591851	AL591851 Mouse DNA
c	996	18	0.8	54147	2	AC121276	AC121276 Mus muscu
c	997	18	0.8	55339	9	AC093593	AC093593 Homo sapi
c	998	18	0.8	55693	2	AC103859	AC103859 Homo sapi
	999	18	0.8	55775	9	AC105426	AC105426 Homo sapi
	1000	18	0.8	56170	2	AC107675	AC107675 Mus muscu

ALIGNMENTS

```
RESULT 1
AY004154
LOCUS       AY004154                2224 bp    mRNA    linear   PRI 25-AUG-2000
DEFINITION  Homo sapiens DEAD box RNA helicase (VASA) mRNA, complete cds.
ACCESSION   AY004154
VERSION     AY004154.1  GI:9438226
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens .
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2224)
  AUTHORS   Castrillon,D.H., Quade,B.J., Wang,T.Y., Quigley,C. and Crum,C.P.
  TITLE     The human VASA gene is specifically expressed in the germ cell
            lineage
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9585-9590 (2000)
  MEDLINE   20402578
   PUBMED   10920202
REFERENCE   2  (bases 1 to 2224)
  AUTHORS   Castrillon,D.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2000) Pathology, Brigham and Women's Hospital, 75
            Francis Street, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..2224
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q11.2-12"
                     /tissue_type="testis"
                     /dev_stage="adult"
     gene            1..2224
                     /gene="VASA"
     CDS             14..2188
                     /gene="VASA"
                     /codon_start=1
                     /product="DEAD box RNA helicase"
                     /protein_id="AAF86585.1"
                     /db_xref="GI:9438227"
                     /translation="MGDEDWEAEINPHMSSYVPIFEKDRYSGENGDNFNRTPASSSEM
                     DDGPSRRDHFMKSGFASGRNFGNRDAGECNKRDNTSTMGGFGVGKSFGNRGFSNSRFE
                     DGDSSGFWRESSNDCEDNPTRNRGFSKRGGYRDGNNSEASGPYRRGGRGSFRGCRGGF
                     GLGSPNNDLDPDECMQRTGGLFGSRRPVLSGTGNGDTSQSRSGSGSERGGYKGLNEEV
                     ITGSGKNSWKSEAEGGESSDTQGPKVTYIPPPPPEDEDSIFAHYQTGINFDKYDTILV
                     EVSGHDAPPAILTFEEANLCQTLNNNIAKAGYTKLTPVQKYSIPIILAGRDLMACAQT
                     GSGKTAAFLLPILAHMMHDGITASRFKELQEPECIIVAPTRELVNQIYLEARKFSFGT
                     CVRAVVIYGGTQLGHSIRQIVQGCNILCATPGRLMDIIGKEKIGLKQIKYLVLDEADR
                     MLDMGFGPEMKKLISCPGMPSKEQRQTLMFSATFPEEIQRLAAEFLKSNYLFVAVGQV
                     GGACRDVQQTVLQVGQFSKREKLVEILRNIGDERTMVFVETKKKADFTATFLCQEKIS
                     TTSIHGDREQREREQALGDFRFGKCPVLVATSVAARGLDIENVQHVINFDLPSTIDEY
                     VHRIGRTGRCGNTGRAISFFDLESDNHLAQPLVKVLTDAQQDVPAWLEEIAFSTYIPG
                     FSGSTRGNVFASVDTRKGKSTLNTAGFSSSRAPNPVDDESWD"
BASE COUNT      678 a    385 c    550 g    611 t
ORIGIN

  Query Match             100.0%;  Score 2172;  DB 9;  Length 2224;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2172;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 73

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 133

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 193

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 253

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 313

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 373
```

```
Qy    361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 433

Qy    421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 493

Qy    481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    494 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 553

Qy    541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    554 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 613

Qy    601 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    614 AGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA 673

Qy    661 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    674 GGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAA 733

Qy    721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    734 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 793

Qy    781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    794 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 853

Qy    841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    854 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 913

Qy    901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    914 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 973

Qy    961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    974 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1033

Qy   1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1034 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTT 1093

Qy   1081 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1094 AAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1153

Qy   1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1154 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1213

Qy   1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1214 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1273

Qy   1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1333

Qy   1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1334 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1393

Qy   1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1394 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1453

Qy   1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1454 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1513

Qy   1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1573

Qy   1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1574 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1633

Qy   1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1634 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1693

Qy   1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1694 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1753

Qy   1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1754 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1813

Qy   1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1814 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1873

Qy   1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1874 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1933

Qy   1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1934 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1993

Qy   1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1994 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATT 2053

Qy   2041 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2054 CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2113

Qy   2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2114 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2173

Qy   2161 GAGTCATGGGAT 2172
          ||||||||||||
Db   2174 GAGTCATGGGAT 2185
```

```
RESULT 2
AF262962
LOCUS       AF262962                2411 bp    mRNA    linear   PRI 30-MAY-2000
DEFINITION  Homo sapiens VASA protein mRNA, complete cds.
ACCESSION   AF262962
VERSION     AF262962.1  GI:8102020
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2411)
  AUTHORS   Rocha,D. and Affara,N.
  TITLE     Cloning and characterization of the human VASA gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2411)
  AUTHORS   Rocha,D. and Affara,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAY-2000) Deparment of Pathology, University of
            Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK
```

```
FEATURES             Location/Qualifiers
     source          1..2411
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
     CDS             18..2192
                     /codon_start=1
                     /product="VASA protein"
                     /protein_id="AAF72705.1"
                     /db_xref="GI:8102021"
                     /translation="MGDEDWEAEINPHMSSYVPIFEKDRYSGENGDNFNRTPASSSEM
                     DDGPSRRDHFMKSGFASGRNFGNRDAGECNKRDNTSTMGGFGVGKSFGNRGFSNSRFE
                     DGDSSGFWRESSNDCEDNPTRNRGFSKRGGYRDGNNSEASGPYRRGGRGSFRGCRGGF
                     GLGSPNNDLDPDECMQRTGGLFGSRRPVLSGTGNGDTSQSRSGSGSERGGYKGLNEEV
                     ITGSGKNSWKSEAEGGESSDTQGPKVTYIPPPPPEDEDSIFAHYQTGINFDKYDTILV
                     EVSGHDAPPAILTFEEANLCQTLNNNIAKAGYTKLTPVQKYSIPIILAGRDLMACAQT
                     GSGKTAAFLLPILAHMMHDGITASRFKELQEPECIIVAPTRELVNQIYLEARKFSFGT
                     CVRAVVIYGGTQLGHSIRQIVQGCNILCATPGRLMDIIGKEKIGLKQIKYLVLDEADR
                     MLDMGFGPEMKKLISCPGMPSKEQRQTLMFSATFPEEIQRLAAEFLKSNYLFVAVGQV
                     GGACRDVQQTVLQVGQFSKREKLVEILRNIGDERTMVFVETKKKADFIATFLCQEKIS
                     TTSIHGDREQREREQALGDFRFGKCPVLVATSVAARGLDIENVQHVINFDLPSTIDEY
                     VHRIGRTGRCGNTGRAISFFDLESDNHLAQPLVKVLTDAQQDVPAWLEEIAFSTYIPG
                     FSGSTRGNVFASVDTRKGKSTLNTAGFSSSQAPNPVDDESWD"
BASE COUNT      771 a    408 c    571 g    661 t
ORIGIN

  Query Match             91.1%;  Score 1978;  DB 9;  Length 2411;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2128;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 83

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 143

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 203

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 263

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 323

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 383

Qy        367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 443

Qy        427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 503

Qy        487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 563

Qy        547 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 623

Qy        607 GGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCT 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 GGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCT 683

Qy        667 GGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCA 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCA 743

Qy        727 AAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTAT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 AAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTAT 803

Qy        787 CAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGAT 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 CAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGAT 863

Qy        847 GCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAAC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAAC 923

Qy        907 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 983

Qy        967 CTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        984 CTTGCAGGACGAGATTTGATGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1043

Qy       1027 CTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 CTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG 1103

Qy       1087 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1104 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1163

Qy       1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1164 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1223

Qy       1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1224 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1283

Qy       1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1284 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1343

Qy       1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1344 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1403

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1463

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1523

Qy       1507 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1583

Qy       1567 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGA 1626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGA 1643

Qy       1627 ACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAA 1686
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       1644 ACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTATTGCAACTTTTCTTTGTCAA 1703

Qy       1687 GAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCT 1746
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1704 GAAAAAATATCAACTACAAGTATTCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCT 1763
```

```
Qy       1747 CTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGA 1806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 CTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGA 1823

Qy       1807 GGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGAT 1866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGAT 1883

Qy       1867 GAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCC 1926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 GAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCC 1943

Qy       1927 TTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGAT 1986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 TTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGAT 2003

Qy       1987 GCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGC 2046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGC 2063

Qy       2047 TTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGC 2106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2064 TTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGC 2123

Qy       2107 ACTTTGAACACAGCTGGGTTTTCTTCTTCAC 2137
              |||||||||||||||||||||||||||||||
Db       2124 ACTTTGAACACAGCTGGGTTTTCTTCTTCAC 2154

RESULT 3
HSM802178
LOCUS       HSM802178               2189 bp    mRNA    linear   PRI 18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp434B1122 (from clone DKFZp434B1122);
            partial cds.
ACCESSION   AL137462
VERSION     AL137462.1  GI:6808043
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2189)
  AUTHORS   Blum,H., Bauersachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by LMU (Ludwig Maximilians University,
            Munich/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp434B1122) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1..2189
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434B1122"
                     /tissue_type="testis"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     gene            1..1910
                     /gene="DKFZp434B1122"
     CDS             <1..1910
                     /gene="DKFZp434B1122"
                     /note="strong similarity to testis specific Helicase (VASA
                     homolog)"
                     /codon_start=3
                     /product="hypothetical protein"
                     /protein_id="CAB70750.1"
                     /db_xref="GI:6808044"
                     /translation="SFGNRGFSNSRFEDGDSSGFWRESSNDCEDNPTRNRGFSKRGGY
                     RDGNNSEASGPYRRGGRGSFRGCRGGFGLGSPNNDLDPDECMQRTGGLFGSRRPVLSG
                     TGNGDTSQSRSGSGSERGGYKGLNEEVITGSGKNSWKSEAEGGESSDTQGPKVTYIPP
                     PPPEDEDSIFAHYQTGINFDKYDTILVEVSGHDAPPAILTFEEANLCQTLNNNIAKAG
                     YTKLTPVQKYSIPIILAGRDLMACAQTGSGKTAAFLLPILAHMMHDGITASRFKELQE
                     PECIIVAPTRELVNQIYLEARKFSFGTCVRAVVIYGGTQLGHSIRQIVQGCNILCATP
                     GRLMDIIGKEKIGLKQIKYLVLDEADRMLDMGFGPEMKKLISCPGMPSKEQRQTLMFS
                     ATFPEEIQRLAAEFLKSNYLFVAVGQVGGACRDVQQTVLQVGQFSKREKLVEILRNIG
                     DERTMVFVETKKKADFIATFLCQEKISTTSIHGDREQREREQALGDFRFGKCPVLVAT
                     SVAARGLDIENVQHVINFDLPSTIDEYVHRIGRTGRCGNTGRAISFFDLESDNHLAQP
                     LVKVLTDAQQDVPAWLEEIAFSTYIPGFSGSTRGNVFASVDTRKGKSTLNTAGFSSSQ
                     APNPVDDESWD"
     polyA_signal    2128..2133
     polyA_site      2147
BASE COUNT      706 a    373 c    505 g    605 t
ORIGIN

  Query Match             79.1%;  Score 1719;  DB 9;  Length 2189;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1869;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        266 AGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTT 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTT 60

Qy        326 TCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCA 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCA 120

Qy        386 AGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTG 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTG 180

Qy        446 GAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACT 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACT 240

Qy        506 TAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTAT 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTAT 300

Qy        566 TAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAG 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAG 360

Qy        626 GTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGT 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGT 420

Qy        686 CAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCC 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCC 480

Qy        746 CTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCG 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCG 540

Qy        806 ACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGA 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGA 600

Qy        866 CTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATA 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATA 660

Qy        926 CTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGA 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGA 720

Qy        986 TGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTC 1045
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        721 TGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTC 780

Qy       1046 ATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTA 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTA 840

Qy       1106 TTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTT 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTT 900

Qy       1166 CTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAA 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAA 960

Qy       1226 TTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATA 1285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATA 1020

Qy       1286 TCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTG 1345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTG 1080

Qy       1346 ATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAA 1405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAA 1140

Qy       1406 TGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTC 1465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTC 1200

Qy       1466 AAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGG 1525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGG 1260

Qy       1526 GTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAG 1585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAG 1320

Qy       1586 AAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAA 1645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAA 1380

Qy       1646 CTAAGAAAAAAGCAGATTTACTGCAACTTTTCTTGTCAAGAAAAAATATCAACTACAA 1705
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       1381 CTAAGAAAAAAGCAGATTTATTGCAACTTTTCTTGTCAAGAAAAAATATCAACTACAA 1440

Qy       1706 GTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTG 1765
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTATTCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTG 1500

Qy       1766 GAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATG 1825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATG 1560

Qy       1826 TGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTG 1885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTG 1620

Qy       1886 GGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGG 1945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGG 1680

Qy       1946 ATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTG 2005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTG 1740

Qy       2006 CATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAG 2065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAG 1800

Qy       2066 GAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGT 2125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGT 1860

Qy       2126 TTTCTTCTTCAC 2137
              ||||||||||||
Db       1861 TTTCTTCTTCAC 1872

RESULT 4
AK093439
LOCUS       AK093439                1984 bp    mRNA    linear   PRI 15-JUL-2002
DEFINITION  Homo sapiens cDNA FLJ36120 fis, clone TESTI2022738, highly similar
            to DEAD BOX PROTEIN 4.
ACCESSION   AK093439
VERSION     AK093439.1  GI:21752310
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens testis cDNA to mRNA, clone_lib:TESTI2
            clone:TESTI2022738.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1984)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1..1984
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TESTI2022738"
                     /tissue_type="testis"
                     /clone_lib="TESTI2"
                     /note="cloning vector: pME18SFL3"
BASE COUNT      612 a    357 c    448 g    567 t
ORIGIN

  Query Match             58.2%;  Score 1265;  DB 9;  Length 1984;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1415;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        720 AGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGC 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 AGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGC 351

Qy        780 ACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGG 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGG 411

Qy        840 ACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAA 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 ACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAA 471

Qy        900 TAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCC 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCC 531
```

```
Qy        960 TATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGC 1019
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        532 TATCATACTTGCAGGACGAGATTTGATGGCTTGCGCTCAAACAGGGTCTGGGAAGACTGC 591

Qy       1020 GGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTT 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 GGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTT 651

Qy       1080 TAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCA 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 TAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCA 711

Qy       1140 GATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATA 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 GATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATA 771

Qy       1200 TGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATG 1259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 TGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATG 831

Qy       1260 TGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGAT 1319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 TGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGAT 891

Qy       1320 CAAATACTTAGTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAAT 1379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 CAAATACTTAGTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAAT 951

Qy       1380 GAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTT 1439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 GAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTT 1011

Qy       1440 CAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTA 1499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 CAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTA 1071

Qy       1500 TCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCT 1559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1072 TCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCT 1131

Qy       1560 CCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGA 1619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1132 CCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGA 1191

Qy       1620 TGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCT 1679
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1192 TGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTATTGCAACTTTTCT 1251

Qy       1680 TTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGA 1739
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1252 TTGTCAAGAAAAAAATATCAACTACAAGTATTCATGGTGATCGGGAACAGAGAGAGCGGGA 1311

Qy       1740 GCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGC 1799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 GCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGC 1371

Qy       1800 TGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTAC 1859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 TGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTAC 1431

Qy       1860 CATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGC 1919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 CATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGC 1491

Qy       1920 AATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATT 1979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1492 AATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATT 1551

Qy       1980 GACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACAT 2039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 GACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACAT 1611

Qy       2040 TCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGG 2099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 TCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGG 1671

Qy       2100 CAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCAC 2137
              ||||||||||||||||||||||||||||||||||||||
Db       1672 CAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCAC 1709
```

```
RESULT 5
AC008914
LOCUS       AC008914               139677 bp    DNA     linear   PRI 19-APR-2002
DEFINITION  Homo sapiens chromosome 5 clone CTD-2270I16, complete sequence.
ACCESSION   AC008914
VERSION     AC008914.7  GI:20198513
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 139677)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 139677)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 139677)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 139677)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 19, 2002 this sequence version replaced gi:19747123.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 0.2.
FEATURES             Location/Qualifiers
     source          1..139677
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2270I16"
BASE COUNT    37913 a  28849 c  28939 g  43976 t
ORIGIN

  Query Match             12.6%;  Score 273;  DB 9;  Length 139677;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-140;
  Matches  273;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1715 GTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCC 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13765 GTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCC 13824

Qy       1775 CAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATG 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13825 CAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATG 13884

Qy       1835 TTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTG 1894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13885 TTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTG 13944

Qy       1895 GTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATT 1954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      13945 GTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATT 14004

Qy       1955 TAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              |||||||||||||||||||||||||||||||||
Db      14005 TAGCACAGCCTCTAGTAAAAGTATTGACAGATG 14037


RESULT 6
AC016632/c
LOCUS       AC016632               176784 bp    DNA     linear   HTG 31-AUG-2001
DEFINITION  Homo sapiens chromosome 5.clone RP11-175M2, WORKING DRAFT SEQUENCE,
            1 unordered piece.
ACCESSION   AC016632
VERSION     AC016632.6  GI:15290348
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 176784)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 176784)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Aug 25, 2001 this sequence version replaced gi:13699595.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 473855
            Center clone name: RPCI-11_175M2
            -------
            Summary Statistics
            Consensus quality: 175566 bases at least Q40
            Consensus quality: 176782 bases at least Q30
            Consensus quality: 176784 bases at least Q20
            Estimated insert size: 182000; pulse field gel estimation
            Estimated insert size: 176784; sum-of-contigs estimation
            Quality coverage:  7.55 in Q20 bases; pulse field gel estimation
            Quality coverage:  7.78 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   176784: contig of 176784 bp in length.
FEATURES             Location/Qualifiers
     source          1..176784
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-175M2"
                     /clone_lib="RPCI human BAC library 11"
BASE COUNT    57392 a  35064 c  33684 g  50644 t
ORIGIN

  Query Match             12.6%;  Score 273;  DB 2;  Length 176784;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-140;
  Matches  273;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1715 GTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTCGCTTTGGAAAGTGCC 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34889 GTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTCGCTTTGGAAAGTGCC 34830
```

```
Qy       1775 CAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATG 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34829 CAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATG 34770

Qy       1835 TTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTG 1894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34769 TTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTG 34710

Qy       1895 GTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATT 1954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34709 GTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATT 34650

Qy       1955 TAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              |||||||||||||||||||||||||||||||||
Db      34649 TAGCACAGCCTCTAGTAAAAGTATTGACAGATG 34617


RESULT 7
AC016639/c
LOCUS       AC016639               182126 bp    DNA     linear   HTG 18-JUL-2000
DEFINITION  Homo sapiens chromosome 5 clone RP11-412L4, WORKING DRAFT SEQUENCE,
            7 ordered pieces.
ACCESSION   AC016639
VERSION     AC016639.5  GI:7711585
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 182126)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182126)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On May 6, 2000 this sequence version replaced gi:7710162.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 564841
            Center clone name: RPCI-11_412L4
            -------
            Summary Statistics
            Consensus quality: 175214 bases at least Q40
            Consensus quality: 180404 bases at least Q30
            Consensus quality: 181112 bases at least Q20
            Estimated insert size: 178000; pulse field gel estimation
            Estimated insert size: 181876; sum-of-contigs estimation
            Quality coverage:  6.41 in Q20 bases; pulse field gel estimation
            Quality coverage:  6.28 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    48575: contig of 48575 bp in length
            *    48576    48675: gap of unknown length
            *    48676    62810: contig of 14135 bp in length
            *    62811    62910: gap of unknown length
            *    62911    81913: contig of 19003 bp in length
            *    81914    82013: gap of unknown length
            *    82014    92166: contig of 10153 bp in length
            *    92167    92266: gap of unknown length
```

```
           *    92267    167478: contig of 75212 bp in length
           *   167479    167578: gap of unknown length
           *   167579    178445: contig of 10867 bp in length
           *   178446    178545: gap of unknown length
           *   178546    182126: contig of 3581 bp in length.
FEATURES             Location/Qualifiers
     source          1. .182126
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-412L4"
                     /clone_lib="RPCI human BAC library 11"
BASE COUNT    58909 a  35906 c  34606 g  52103 t    602 others
ORIGIN

  Query Match             12.6%;  Score 273;  DB 2;  Length 182126;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-140;
  Matches  273;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1715 GTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCC 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      36792 GTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCC 36733

Qy       1775 CAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATG 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      36732 CAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATG 36673

Qy       1835 TTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTG 1894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      36672 TTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTG 36613

Qy       1895 GTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATT 1954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      36612 GTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATT 36553

Qy       1955 TAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              |||||||||||||||||||||||||||||||||
Db      36552 TAGCACAGCCTCTAGTAAAAGTATTGACAGATG 36520


RESULT 8
AC022265
LOCUS       AC022265              204250 bp    DNA     linear   HTG 13-AUG-2000
DEFINITION  Homo sapiens chromosome 5 clone RP11-332C17, WORKING DRAFT
            SEQUENCE, 16 unordered pieces.
ACCESSION   AC022265
VERSION     AC022265.3  GI:9719812
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 204250)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204250)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Aug 7, 2000 this sequence version replaced gi:7940395.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0332C17
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 87%
            Sequencing vector: plasmid; 13%
            Chemistry: Dye-primer ET; 85% of reads
            Chemistry: Dye-terminator Big Dye; 15% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 196499 bases at least Q40
            Consensus quality: 199226 bases at least Q30
            Consensus quality: 200804 bases at least Q20
            Insert size: 205000; agarose-fp
            Insert size: 202750; sum-of-contigs
            Quality coverage: 4.50 in Q20 bases; agarose-fp
            Quality coverage: 4.65 in Q20 bases; sum-of-contigs.
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1062: contig of 1062 bp in length
            *     1063     1162: gap of unknown length
            *     1163     2536: contig of 1374 bp in length
            *     2537     2636: gap of unknown length
            *     2637     5791: contig of 3155 bp in length
            *     5792     5891: gap of unknown length
            *     5892    10946: contig of 5055 bp in length
            *    10947    11046: gap of unknown length
            *    11047    15725: contig of 4679 bp in length
            *    15726    15825: gap of unknown length
            *    15826    21169: contig of 5344 bp in length
            *    21170    21269: gap of unknown length
            *    21270    27800: contig of 6531 bp in length
            *    27801    27900: gap of unknown length
            *    27901    33276: contig of 5376 bp in length
            *    33277    33376: gap of unknown length
            *    33377    42711: contig of 9335 bp in length
            *    42712    42811: gap of unknown length
            *    42812    56046: contig of 13235 bp in length
            *    56047    56146: gap of unknown length
            *    56147    77263: contig of 21117 bp in length
            *    77264    77363: gap of unknown length
            *    77364    97974: contig of 20611 bp in length
            *    97975    98074: gap of unknown length
            *    98075   122272: contig of 24198 bp in length
            *   122273   122372: gap of unknown length
            *   122373   148199: contig of 25827 bp in length
            *   148200   148299: gap of unknown length
            *   148300   175035: contig of 26736 bp in length
            *   175036   175135: gap of unknown length
            *   175136   204250: contig of 29115 bp in length.
FEATURES             Location/Qualifiers
     source          1. .204250
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-332C17"
     misc_feature    1. .1062
                     /note="assembly_name:Contig12"
     misc_feature    1163. .2536
                     /note="assembly_name:Contig13"
     misc_feature    2637. .5791
                     /note="assembly_name:Contig14"
     misc_feature    5892. .10946
                     /note="assembly_name:Contig15"
     misc_feature    11047. .15725
                     /note="assembly_name:Contig16"
     misc_feature    15826. .21169
                     /note="assembly_name:Contig17"
     misc_feature    21270. .27800
                     /note="assembly_name:Contig18"
     misc_feature    27901. .33276
                     /note="assembly_name:Contig19"
     misc_feature    33377. .42711
```

```
                 /note="assembly_name:Contig20
                 clone_end:T7
                 vector_side:right"
     misc_feature    42812..56046
                 /note="assembly_name:Contig21"
     misc_feature    56147..77263
                 /note="assembly_name:Contig22"
     misc_feature    77364..97974
                 /note="assembly_name:Contig23"
     misc_feature    98075..122272
                 /note="assembly_name:Contig24"
     misc_feature    122373..148199
                 /note="assembly_name:Contig25"
     misc_feature    148300..175035
                 /note="assembly_name:Contig26"
     misc_feature    175136..204250
                 /note="assembly_name:Contig27"
BASE COUNT    62979 a  41247 c  40079 g  58442 t   1503 others
ORIGIN

  Query Match             12.6%;  Score 273;  DB 2;  Length 204250;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-140;
  Matches  273;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1715 GTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCC 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21580 GTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCC 21639

Qy       1775 CAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATG 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21640 CAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATG 21699

Qy       1835 TTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTG 1894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21700 TTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTG 21759

Qy       1895 GTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATT 1954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21760 GTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATT 21819

Qy       1955 TAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
              |||||||||||||||||||||||||||||||||
Db      21820 TAGCACAGCCTCTAGTAAAAGTATTGACAGATG 21852


RESULT 9
AC022265/c
LOCUS       AC022265              204250 bp    DNA     linear   HTG 13-AUG-2000
DEFINITION  Homo sapiens chromosome 5 clone RP11-332C17, WORKING DRAFT
            SEQUENCE, 16 unordered pieces.
ACCESSION   AC022265
VERSION     AC022265.3  GI:9719812
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 204250)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204250)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Aug 7, 2000 this sequence version replaced gi:7940395.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ------------- Project Information -------------
            Center project name: H_NH0332C17
            ------------- Summary Statistics -------------
            Sequencing vector: M13; 87%
            Sequencing vector: plasmid; 13%
            Chemistry: Dye-primer ET; 85% of reads
            Chemistry: Dye-terminator Big Dye; 15% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 196499 bases at least Q40
            Consensus quality: 199226 bases at least Q30
            Consensus quality: 200804 bases at least Q20
            Insert size: 205000; agarose-fp
            Insert size: 202750; sum-of-contigs
            Quality coverage: 4.50 in Q20 bases; agarose-fp
            Quality coverage: 4.65 in Q20 bases; sum-of-contigs
            ----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1      1062: contig of 1062 bp in length
            *       1063      1162: gap of unknown length
            *       1163      2536: contig of 1374 bp in length
            *       2537      2636: gap of unknown length
            *       2637      5791: contig of 3155 bp in length
            *       5792      5891: gap of unknown length
            *       5892     10946: contig of 5055 bp in length
            *      10947     11046: gap of unknown length
            *      11047     15725: contig of 4679 bp in length
            *      15726     15825: gap of unknown length
            *      15826     21169: contig of 5344 bp in length
            *      21170     21269: gap of unknown length
            *      21270     27800: contig of 6531 bp in length
            *      27801     27900: gap of unknown length
            *      27901     33276: contig of 5376 bp in length
            *      33277     33376: gap of unknown length
            *      33377     42711: contig of 9335 bp in length
            *      42712     42811: gap of unknown length
            *      42812     56046: contig of 13235 bp in length
            *      56047     56146: gap of unknown length
            *      56147     77263: contig of 21117 bp in length
            *      77264     77363: gap of unknown length
            *      77364     97974: contig of 20611 bp in length
            *      97975     98074: gap of unknown length
            *      98075    122272: contig of 24198 bp in length
            *     122273    122372: gap of unknown length
            *     122373    148199: contig of 25827 bp in length
            *     148200    148299: gap of unknown length
            *     148300    175035: contig of 26736 bp in length
            *     175036    175135: gap of unknown length
            *     175136    204250: contig of 29115 bp in length.
FEATURES             Location/Qualifiers
     source          1..204250
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-332C17"
     misc_feature    1..1062
                     /note="assembly_name:Contig12"
     misc_feature    1163..2536
                     /note="assembly_name:Contig13"
     misc_feature    2637..5791
                     /note="assembly_name:Contig14"
     misc_feature    5892..10946
                     /note="assembly_name:Contig15"
     misc_feature    11047..15725
                     /note="assembly_name:Contig16"
     misc_feature    15826..21169
                     /note="assembly_name:Contig17"
```

```
    misc_feature    21270..27800
                    /note="assembly_name:Contig18"
    misc_feature    27901..33276
                    /note="assembly_name:Contig19"
    misc_feature    33377..42711
                    /note="assembly_name:Contig20
                    clone_end:T7
                    vector_side:right"
    misc_feature    42812..56046
                    /note="assembly_name:Contig21"
    misc_feature    56147..77263
                    /note="assembly_name:Contig22"
    misc_feature    77364..97974
                    /note="assembly_name:Contig23"
    misc_feature    98075..122272
                    /note="assembly_name:Contig24"
    misc_feature    122373..148199
                    /note="assembly_name:Contig25"
    misc_feature    148300..175035
                    /note="assembly_name:Contig26"
    misc_feature    175136..204250
                    /note="assembly_name:Contig27"
BASE COUNT    62979 a  41247 c  40079 g  58442 t   1503 others
ORIGIN

  Query Match              8.0%;  Score 174;  DB 2;  Length 204250;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-85;
  Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1299 AAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGA 1358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170192 AAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGA 170133

Qy       1359 TATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGA 1418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170132 TATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGA 170073

Qy       1419 ACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTT 1472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170072 ACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTT 170019


RESULT 10
AY100475
LOCUS       AY100475                 524 bp    DNA     linear   MAM 15-JUN-2002
DEFINITION  Equus caballus VASA-like protein gene, partial sequence.
ACCESSION   AY100475
VERSION     AY100475.1  GI:21429223
KEYWORDS    .
SOURCE      horse.
  ORGANISM  Equus caballus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE   1  (bases 1 to 524)
  AUTHORS   Woods,B.G., Ginther,O.J., Wentworth,A., Wentworth,B. and
            Wiltbank,M.
  TITLE     Equine VASA Homolog
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 524)
  AUTHORS   Woods,B.G., Ginther,O.J., Wentworth,A., Wentworth,B. and
            Wiltbank,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2002) Animal Health and Biomedical Sciences,
            University of Wisconsin-Madison, 1656 Linden Drive, Madison, WI
            53706, USA
FEATURES             Location/Qualifiers
     source          1..524
                     /organism="Equus caballus"
                     /db_xref="taxon:9796"
    misc_feature     <1..>524
                     /note="similar to Drosophila melanogaster VASA protein"
BASE COUNT      153 a     94 c    115 g    151 t     11 others
ORIGIN

  Query Match              2.7%;  Score 58;  DB 4;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-20;
  Matches   58;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1984 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTC 2041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTC 512


RESULT 11
S75275
LOCUS       S75275                  3030 bp    mRNA    linear   ROD 15-MAY-1995
DEFINITION  RVLG=vasa-like gene protein [rats, Wistar-Imanishi, testis, mRNA,
            3030 nt].
ACCESSION   S75275
VERSION     S75275.1  GI:806463
KEYWORDS    .
SOURCE      Rattus sp. testis Wistar-Imanishi.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3030)
  AUTHORS   Komiya,T. and Tanigawa,Y.
  TITLE     Cloning of a gene of the DEAD box protein family which is
            specifically expressed in germ cells in rats
  JOURNAL   Biochem. Biophys. Res. Commun. 207 (1), 405-410 (1995)
  MEDLINE   95160706
   PUBMED   7857296
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 159937] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1..3030
                     /organism="Rattus sp."
                     /db_xref="taxon:10118"
     gene            1..3030
                     /gene="RVLG"
                     /note="vasa-like gene protein, RVLG protein"
     CDS             266..2407
                     /gene="RVLG"
                     /note="putative DEAD box protein; This sequence comes from
                     Fig. 1; RVLG protein"
                     /codon_start=1
                     /product="vasa-like gene protein"
                     /protein_id="AAB33364.1"
                     /db_xref="GI:806464"
                     /translation="MGDEDWEAEILKPHVSSYVPVFEKDKYSSGANGDTFNRTSASSS
                     EMEDGPSGRDHFMRSGFSSGRNLGNRDIGESSKRETTSTTGGFGRGKGFGNRGFLNNK
                     FEEGDSSGFWKESTNDCEDTQTRSRGFSKRGGYPDGNDSEASGPFRRGGRDSEYDQDQ
                     GSQRGGGLFGSRKPAASDSGSGDTFQSRSGNARGAYKGLNEEVVTGSGKNSWKSEAEG
                     GESSDIQGPKVTYIPPPPPEDEDSIFAHYQTGINFDKYDTILVEVSGHDAPPAILTFE
                     EANLCQTLNNNIAKAGYTKLTPVQKYSIPIVLAGRDLMACAQTGSGKTAAFLLPILAH
                     MMRDGITASRFKELQEPECIIVAPTRELINQIYLEARKFSFGTCVRAVVIYGGTQFGH
                     SIRQIVQGCNILCATPGRLMDIIGKEKIGLKQVKYLVLDEADRMLDMGFGPEMKKLIS
                     CPGMPSKEQRQTLLFSATFPEEIQRLAGEFLKSNYLFVAVGQVGGACRDVQQSILQVG
                     PVFKKRKLVEILRNIGDERPMVFVETKKKADFIATFLCQEKISTTSIHGDREQREREQ
                     ALGDFRCGKCPVLVATSVAARGLDIENVQHVINFNLPSTIDEYVHRIGRTGRCGNTGR
                     AISFFDTESDNHLAQPLVKVLSDAQQDVPAWLEEIAFSSYAPPSFSNSTRGAVFASFD
                     TRKNFQGKNTLNTAGISSAQAPNPVDDESWD"
BASE COUNT      900 a    551 c    723 g    856 t
ORIGIN

  Query Match              2.4%;  Score 53;  DB 10;  Length 3030;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-17;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGT 1193
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1361 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGT 1413
```

```
RESULT 12
MUSDVH
LOCUS       MUSDVH                  1930 bp    mRNA    linear   ROD 04-FEB-1999
DEFINITION  Mouse mRNA for drosophila vasa homologue, partial cds.
ACCESSION   D14859
VERSION     D14859.1  GI:286074
KEYWORDS    RNA helicase; drosophila vasa homologue.
SOURCE      Mus musculus (strain BALB/c) (library: lambda gt10) adult and
            embryo gonad and testis primordial germ cell, spermatogonium and
            spermatocyte cDNA to mRNA, clone mVH.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (sites)
  AUTHORS   Fujiwara,Y., Komiya,T., Kawabata,H., Sato,M., Fujimoto,H.,
            Furusawa,M. and Noce,T.
  TITLE     Isolation of a DEAD-family protein gene that encodes a murine
            homolog of Drosophila vasa and its specific expression in germ cell
            lineage
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 91 (25), 12258-12262 (1994)
  MEDLINE   95083681
REFERENCE   2  (bases 1 to 1930)
  AUTHORS   Noce,T.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1930)
  AUTHORS   Noce,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-1993) Toshiaki Noce, Mitsubishi Kasei Institute
            of Life Sciences, Developmental Biology; 11 Minamiooya, Machida,
            Tokyo 194, Japan (Tel:0427-24-6246, Fax:0427-29-1252)
FEATURES             Location/Qualifiers
     source          1..1930
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /cell_type="primordial germ cell, spermatogonium,
                     spermatocyte"
                     /tissue_type="gonad, testis"
                     /clone_lib="lambda gt10"
                     /dev_stage="adult, embryo"
     CDS             <1..1914
                     /function="RNA helicase"
                     /codon_start=1
                     /product="Drosophila vasa homologue"
                     /protein_id="BAA03584.1"
                     /db_xref="GI:286075"
                     /translation="FGRGKGFGNRGFLNNKFEEGDSSGFWKESNNDCEDNQTRSRGFS
                     KRGGCQDGNDSEASGPFRRGGRGSFRGCRGGFGLGRPNSESDQDQGTQCGGGFLVLGK
                     PAASDSGNGDTYQSRSGSGRGGYKGLNEEVVTGSGKNSWKSETEGGESSDSQGPKVTY
                     IPPPPPEDEDSIFAHYQTGINFDKYDTILVEVSGHDAPPAILTFEEANLCQTLNNNIR
                     KAGYTKLTPVQKYTIPIVLAGRDLMACAQTGSGKTAAFLLPILAHMMRDGITASRFKE
                     LQEPECIIVAPTRELINQIYLEARKFSFGTCVISVVIYGGTQFGHSVRQIVQGCNILC
                     ATPGRLMDIIGKEKIGLKQVKYLVLDEADSMLDMGFAPEIKKLISCPGMPSKEQHQTL
                     LFSATFPEEIQRLAGDFLKSNYLFVAVGQVGGACRDVQQTILQVGQYQKEKSLLRFYE
                     NIGDERTMVFVETKKKADFIATFLCQEKISSTSIHGDREQREREQALGDFRCGKCPVL
                     VATSVAARGLDIENVQHVINFDLPSTIDEYVHRIGRTGRCGNTGRAISFFDTDSDNHL
                     AQPLVKVLSDAQQDVPAWLEEIAFSTYVPPSFSSSTRGGAVFASVDTRKNYQGKAHVE
                     YSGDFFFTSSQSS"
BASE COUNT      577 a    356 c    487 g    510 t
ORIGIN

  Query Match             2.3%;  Score 50;  DB 10;  Length 1930;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGC 1019
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGC 761


RESULT 13
AC124894/c
LOCUS       AC124894              108795 bp    DNA     linear   HTG 03-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-5A4, *** SEQUENCING IN PROGRESS ***,
            56 unordered pieces.
ACCESSION   AC124894
VERSION     AC124894.1  GI:21465340
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 108795)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 108795)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 108795)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GBAA
```

```
    Center clone name: CH230-5A4
 -------------- Summary Statistics
    Sequencing vector: Plasmid;
    Chemistry: Dye-terminator Big Dye: 100% of reads
    Assembly program: Phrap; version 0.990329
    Consensus quality: 56701 bases at least Q40
    Consensus quality: 62823 bases at least Q30
    Consensus quality: 66969 bases at least Q20
 --------------
 * NOTE: Estimated insert size may differ from sequence length
 *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 56 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *        1     1168: contig of 1168 bp in length
 *     1169     1268: gap of unknown length
 *     1269     2361: contig of 1093 bp in length
 *     2362     2461: gap of unknown length
 *     2462     3595: contig of 1134 bp in length
 *     3596     3695: gap of unknown length
 *     3696     5046: contig of 1351 bp in length
 *     5047     5146: gap of unknown length
 *     5147     6256: contig of 1110 bp in length
 *     6257     6356: gap of unknown length
 *     6357     7858: contig of 1502 bp in length
 *     7859     7958: gap of unknown length
 *     7959     9544: contig of 1586 bp in length
 *     9545     9644: gap of unknown length
 *     9645    10848: contig of 1204 bp in length
 *    10849    10948: gap of unknown length
 *    10949    12755: contig of 1807 bp in length
 *    12756    12855: gap of unknown length
 *    12856    14295: contig of 1440 bp in length
 *    14296    14395: gap of unknown length
 *    14396    15508: contig of 1113 bp in length
 *    15509    15608: gap of unknown length
 *    15609    16792: contig of 1184 bp in length
 *    16793    16892: gap of unknown length
 *    16893    18735: contig of 1843 bp in length
 *    18736    18835: gap of unknown length
 *    18836    20186: contig of 1351 bp in length
 *    20187    20286: gap of unknown length
 *    20287    21888: contig of 1602 bp in length
 *    21889    21988: gap of unknown length
 *    21989    23783: contig of 1795 bp in length
 *    23784    23883: gap of unknown length
 *    23884    25610: contig of 1727 bp in length
 *    25611    25710: gap of unknown length
 *    25711    27880: contig of 2170 bp in length
 *    27881    27980: gap of unknown length
 *    27981    29060: contig of 1080 bp in length
 *    29061    29160: gap of unknown length
 *    29161    30586: contig of 1426 bp in length
 *    30587    30686: gap of unknown length
 *    30687    32406: contig of 1720 bp in length
 *    32407    32506: gap of unknown length
 *    32507    34290: contig of 1784 bp in length
 *    34291    34390: gap of unknown length
 *    34391    36703: contig of 2313 bp in length
 *    36704    36803: gap of unknown length
 *    36804    37830: contig of 1027 bp in length
 *    37831    37930: gap of unknown length
 *    37931    39017: contig of 1087 bp in length
 *    39018    39117: gap of unknown length
 *    39118    40137: contig of 1020 bp in length
 *    40138    40237: gap of unknown length
 *    40238    42389: contig of 2152 bp in length
 *    42390    42489: gap of unknown length
 *    42490    43589: contig of 1100 bp in length
 *    43590    43689: gap of unknown length
 *    43690    45802: contig of 2113 bp in length
 *    45803    45902: gap of unknown length
 *    45903    48014: contig of 2112 bp in length
 *    48015    48114: gap of unknown length
 *    48115    49732: contig of 1618 bp in length
 *    49733    49832: gap of unknown length
 *    49833    51415: contig of 1583 bp in length
 *    51416    51515: gap of unknown length
 *    51516    52858: contig of 1343 bp in length
 *    52859    52958: gap of unknown length
 *    52959    55075: contig of 2117 bp in length
 *    55076    55175: gap of unknown length
 *    55176    57312: contig of 2137 bp in length
 *    57313    57412: gap of unknown length
 *    57413    59494: contig of 2082 bp in length
 *    59495    59594: gap of unknown length
 *    59595    61146: contig of 1552 bp in length
 *    61147    61246: gap of unknown length
 *    61247    63268: contig of 2022 bp in length
 *    63269    63368: gap of unknown length
 *    63369    65294: contig of 1926 bp in length
 *    65295    65394: gap of unknown length
 *    65395    67457: contig of 2063 bp in length
 *    67458    67557: gap of unknown length
 *    67558    69843: contig of 2286 bp in length
 *    69844    69943: gap of unknown length
 *    69944    72337: contig of 2394 bp in length
 *    72338    72437: gap of unknown length
 *    72438    73987: contig of 1550 bp in length
 *    73988    74087: gap of unknown length
 *    74088    76226: contig of 2139 bp in length
 *    76227    76326: gap of unknown length
 *    76327    79288: contig of 2962 bp in length
 *    79289    79388: gap of unknown length
 *    79389    81977: contig of 2589 bp in length
 *    81978    82077: gap of unknown length
 *    82078    84162: contig of 2085 bp in length
 *    84163    84262: gap of unknown length
 *    84263    87197: contig of 2935 bp in length
 *    87198    87297: gap of unknown length
 *    87298    90153: contig of 2856 bp in length
 *    90154    90253: gap of unknown length
 *    90254    92307: contig of 2054 bp in length
 *    92308    92407: gap of unknown length
 *    92408    94844: contig of 2437 bp in length
 *    94845    94944: gap of unknown length
 *    94945    97858: contig of 2914 bp in length
 *    97859    97958: gap of unknown length
 *    97959   100200: contig of 2242 bp in length
 *   100201   100300: gap of unknown length
 *   100301   102763: contig of 2463 bp in length
 *   102764   102863: gap of unknown length

  Query Match             2.2%;  Score 47;  DB 2;  Length 108795;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-14;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        757 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTT 803
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      89719 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTT 89673


RESULT 14
AC105884/c
LOCUS       AC105884               114024 bp    DNA     linear   HTG 13-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-13O2, *** SEQUENCING IN PROGRESS ***,
            52 unordered pieces.
ACCESSION   AC105884
VERSION     AC105884.2  GI:21736908
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
```

```
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 114024)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 114024)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 114024)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 12, 2002 this sequence version replaced gi:18104791.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GNZK
                Center clone name: CH230-1302
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 62592 bases at least Q40
                Consensus quality: 66494 bases at least Q30
                Consensus quality: 69970 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 52 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      1255: contig of 1255 bp in length
            *     1256      1355: gap of unknown length
            *     1356      2912: contig of 1557 bp in length
            *     2913      3012: gap of unknown length
            *     3013      4498: contig of 1486 bp in length
            *     4499      4598: gap of unknown length
            *     4599      5889: contig of 1291 bp in length
            *     5890      5989: gap of unknown length
            *     5990      7578: contig of 1589 bp in length
            *     7579      7678: gap of unknown length
            *     7679      9026: contig of 1348 bp in length
            *     9027      9126: gap of unknown length
            *     9127     10186: contig of 1060 bp in length
            *    10187     10286: gap of unknown length
            *    10287     11373: contig of 1087 bp in length
            *    11374     11473: gap of unknown length
            *    11474     13029: contig of 1556 bp in length
            *    13030     13129: gap of unknown length
            *    13130     14667: contig of 1538 bp in length
            *    14668     14767: gap of unknown length
            *    14768     16235: contig of 1468 bp in length
            *    16236     16335: gap of unknown length
            *    16336     17357: contig of 1022 bp in length
            *    17358     17457: gap of unknown length
            *    17458     18859: contig of 1402 bp in length
            *    18860     18959: gap of unknown length
            *    18960     20678: contig of 1719 bp in length
            *    20679     20778: gap of unknown length
            *    20779     21924: contig of 1146 bp in length
            *    21925     22024: gap of unknown length
            *    22025     23056: contig of 1032 bp in length
            *    23057     23156: gap of unknown length
            *    23157     24351: contig of 1195 bp in length
            *    24352     24451: gap of unknown length
            *    24452     25662: contig of 1211 bp in length
            *    25663     25762: gap of unknown length
            *    25763     27168: contig of 1406 bp in length
            *    27169     27268: gap of unknown length
            *    27269     28375: contig of 1107 bp in length
            *    28376     28475: gap of unknown length
            *    28476     30236: contig of 1761 bp in length
            *    30237     30336: gap of unknown length
            *    30337     31660: contig of 1324 bp in length
            *    31661     31760: gap of unknown length
            *    31761     33107: contig of 1347 bp in length
            *    33108     33207: gap of unknown length
            *    33208     35825: contig of 2618 bp in length
            *    35826     35925: gap of unknown length
            *    35926     37404: contig of 1479 bp in length
            *    37405     37504: gap of unknown length
            *    37505     39269: contig of 1765 bp in length
            *    39270     39369: gap of unknown length
            *    39370     41534: contig of 2165 bp in length
            *    41535     41634: gap of unknown length
            *    41635     43062: contig of 1428 bp in length
            *    43063     43162: gap of unknown length
            *    43163     45218: contig of 2056 bp in length
            *    45219     45318: gap of unknown length
            *    45319     46889: contig of 1571 bp in length
            *    46890     46989: gap of unknown length
```

```
*    46990    48117: contig of 1128 bp in length
*    48118    48217: gap of unknown length
*    48218    50706: contig of 2489 bp in length
*    50707    50806: gap of unknown length
*    50807    52398: contig of 1592 bp in length
*    52399    52498: gap of unknown length
*    52499    54416: contig of 1918 bp in length
*    54417    54516: gap of unknown length
*    54517    56139: contig of 1623 bp in length
*    56140    56239: gap of unknown length
*    56240    58833: contig of 2594 bp in length
*    58834    58933: gap of unknown length
*    58934    61254: contig of 2321 bp in length
*    61255    61354: gap of unknown length
*    61355    63634: contig of 2280 bp in length
*    63635    63734: gap of unknown length
*    63735    66620: contig of 2886 bp in length
*    66621    66720: gap of unknown length
*    66721    69735: contig of 3015 bp in length
*    69736    69835: gap of unknown length
*    69836    73021: contig of 3186 bp in length
*    73022    73121: gap of unknown length
*    73122    75151: contig of 2030 bp in length
*    75152    75251: gap of unknown length
*    75252    77299: contig of 2048 bp in length
*    77300    77399: gap of unknown length
*    77400    80847: contig of 3448 bp in length
*    80848    80947: gap of unknown length
*    80948    82958: contig of 2011 bp in length
*    82959    83058: gap of unknown length
*    83059    86408: contig of 3350 bp in length
*    86409    86508: gap of unknown length
*    86509    89930: contig of 3422 bp in length
*    89931    90030: gap of unknown length
*    90031    95181: contig of 5151 bp in length
*    95182    95281: gap of unknown length
*    95282   100060: contig of 4779 bp in length
*   100061   100160: gap of unknown length
*   100161   103756: contig of 3596 bp in length
*   103757   103856: gap of unknown length
*   103857   108054: contig of 4198 bp in length
*   108055   108154: gap of unknown length
*   108155   114024: contig of 5870 bp in length.
FEATURES             Location/Qualifiers
     source          1. .114024
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"

  Query Match             2.2%;  Score 47;  DB 2;  Length 114024;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-14;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        757 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTT 803
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      10537 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTT 10491


RESULT 15
AE013856/c
LOCUS       AE013856            14350 bp    DNA     linear   BCT 26-JUL-2002
DEFINITION  Yersinia pestis KIM section 256 of 415 of the complete genome.
ACCESSION   AE013856 AE009952
VERSION     AE013856.1  GI:21959401
KEYWORDS    .
SOURCE      Yersinia pestis KIM.
  ORGANISM  Yersinia pestis KIM
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Yersinia.
REFERENCE   1  (bases 1 to 14350)
  AUTHORS   Deng,W., Burland,V., Plunkett,G. III, Boutin,A., Mayhew,G.F.,
            Liss,P., Perna,N.T., Rose,D.J., Mau,B., Zhou,S., Schwartz,D.C.,
            Fetherston,J.D., Lindler,L.E., Brubaker,R.R., Plana,G.V.,
            Straley,S.C., McDonough,K.A., Nilles,M.L., Matson,J.S.,
            Blattner,F.R. and Perry,R.D.
  TITLE     Genome Sequence of Yersinia pestis KIM
  JOURNAL   J. Bacteriol. 184 (16), 4601-4611 (2002)
   PUBMED   12142430
REFERENCE   2  (bases 1 to 14350)
  AUTHORS   Deng,W., Burland,V., Plunkett,G. III, Boutin,A., Mayhew,G.F.,
            Liss,P., Perna,N.T., Rose,D.J., Mau,B., Zhou,S., Schwartz,D.C.,
            Fetherston,J.D., Lindler,L.E., Brubaker,R.R., Plana,G.V.,
            Straley,S.C., McDonough,K.A., Nilles,M.L., Matson,J.S.,
            Blattner,F.R. and Perry,R.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445
            Henry Mall, Madison, WI 53706, USA
FEATURES             Location/Qualifiers
     source          1. .14350
                     /organism="Yersinia pestis KIM"
                     /strain="KIM"
                     /db_xref="taxon:187410"
     gene            complement(78. .1460)
                     /gene="dbpA"
                     /note="y2531"
     CDS             complement(78. .1460)
                     /gene="dbpA"
                     /function="enzyme; RNA synthesis, modification, DNA
                     transcription"
                     /note="residues 4 to 460 of 460 are 68.84 pct identical to
                     residues 2 to 457 of 457 from E. coli K12 : B1343;
                     residues 4 to 460 of 460 are 68.62 pct identical to
                     residues 2 to 457 of 457 from GenPept :
                     >gb|AAG56454.1|AE005371_10 (AE005371) ATP-dependent RNA
                     helicase [Escherichia coli O157:H7 EDL933]"
                     /codon_start=1
                     /transl_table=11
                     /product="ATP-dependent RNA helicase"
                     /protein_id="AAM86087.1"
                     /db_xref="GI:21959402"
                     /translation="MSTTSFSSLTLPAEQLSNLNELGYTEMTPVQAAVLPAILNGQDV
                     RAKAKTGSGKTAAFGIGLLDKIAVGEFVTQALVLCPTRELADQVSKELRRLARFTQNI
                     KILTLCGGQPMGHQLDSLVHAPHIVVGTPGRIQEHLRKKTLVLDELKTLVLDEADRML
                     DMGFSDAIDDVIAYTPPQRQTLLFSATYPSGIEQISARVQRQPLNVEVGDTDEEPAIE
                     QVFFETTREKRLPLLISVLSHYQPASCVVFCNTKKDCQSVYESLESRGISVLALHGDL
                     EQRDRDQVLVRFANRSCRVLVATDVAARGLDIKDLELVINFELAFDPEIHIHRIGRTG
                     RAGMSGLAVSLCTPQEMVRAHTIEDYLQIKLKWTPAEQVSRSTNIMLEPEMVTLCIDG
                     GRKAKIRPGDILGALTGDAGLTAADVGKIDMFPVHAYVAIRKASAKRALQQLQQGKIK
                     GKNCKARLLK"
     gene            1773. .2954
                     /gene="purT"
                     /note="y2533"
     CDS             1773. .2954
                     /gene="purT"
                     /function="enzyme; purine ribonucleotide biosynthesis"
                     /note="residues 1 to 392 of 393 are 78.31 pct identical to
                     residues 1 to 392 of 392 from E. coli K12 : B1849"
                     /codon_start=1
                     /transl_table=11
                     /product="phosphoribosylglycinamide formyltransferase 2"
                     /protein_id="AAM86088.1"
                     /db_xref="GI:21959403"
                     /translation="MLTIGTALRPGATRVMLLGAGELGKEVAIECQRLGLEVIAVDRY
                     ADAPAMHVAHRSHVINMLDGAALKQLVAQEKPHYIVPEIEAIATDMLVELEKMGQHVV
                     PCAEATRLTMNREGIRRLAAETLQLPTSSYRFADTDSAFFQAVRDIGYPCIVKPVMSS
                     SGKGQSLIRSEEHLQAAWEYAQQGGRAGSGRVIIEGLVHFDFEITLLTIRAVDGIHFC
                     APIGHRQEDGDYRESWQPQAMSDIALQRAKEISAQVVTALGGFGLFGVELFVCGDDVI
                     FSEVSPRPHDTGMVTLISQNMSEFALHVRAFLGLPIGTIRQYGAAASAVILPELTSQN
                     ITYRGLETALIGDTQIRLFGKPEIAGKRRLGVALAVADNIETAIEVAKKAAGNIEVSG
                     E"
     gene            complement(2151. .2369)
                     /gene="y2532"
     CDS             complement(2151. .2369)
                     /gene="y2532"
                     /function="unknown"
                     /note="residues 41 to 68 of 72 are 46.42 pct identical to
                     residues 841 to 868 of 1147 from GenPept :
```

```
                     >emb|CAB40374.1| (AJ225088) Starch synthase isoform SS III
                     [Vigna unguiculata]"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical"
                     /protein_id="AAM86089.1"
                     /db_xref="GI:21959404"
                     /translation="MNQTFDNHAPATGTATLLCIFPSSLKVFLAADQALAFTRRTHYR
                     FNNARVTYVAHGLKEGAICVGKAVTGRG"
     gene            complement(3080. .3337)
                     /gene="y2534"
     CDS             complement(3080. .3337)
                     /gene="y2534"
                     /note="residues 1 to 59 of 85 are 59.32 pct identical to
                     residues 1 to 58 of 59 from E. coli K12 : B1811; residues
                     1 to 59 of 85 are 62.71 pct identical to residues 1 to 58
                     of 59 from GenPept : >gb|AAL20738.1| (AE008781) putative
                     cytoplasmic protein [Salmonella typhimurium LT2]"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAM86090.1"
                     /db_xref="GI:21959405"
                     /translation="MLAGMPSLSHEEQQEAVERIHKFMSEGMSSGEAIALVAAEIRER
                     HQNDPQAMAIFEDHDFDEHTESDYRRDDEPDADDIEDPYEG"
     gene            3502. .4878
                     /gene="pabB"
                     /note="y2535"
     CDS             3502. .4878
                     /gene="pabB"
                     /function="enzyme; biosynthesis of cofactors, carriers:
                     Folic acid"
                     /note="residues 11 to 452 of 458 are 66.81 pct identical
                     to residues 11 to 451 of 453 from E. coli K12 : B1812;
                     residues 7 to 452 of 458 are 67.33 pct identical to
                     residues 8 to 452 of 454 from GenPept : >gb|AAL20739.1|
                     (AE008781) p-aminobenzoate synthetase, component I
                     [Salmonella typhimurium LT2]"
                     /codon_start=1
                     /transl_table=11
                     /product="p-aminobenzoate synthetase, component I"
                     /protein_id="AAM86091.1"
                     /db_xref="GI:21959406"
                     /translation="MCEKFLTVKELPYHPDALLQQFTPLANQAWAMLLHSGFAQHPHS
                     RFDILVADPKITLTTHGDKTVIVSDQGQETSLADPFSLLQQQLDKFAPAMPAHPDLPF
                     QGGALGLFGYDLGRRVEKLPELAAHDIALPDMAIGLYDWALIADHSRQKLTLVSLGDA
                     EQRLHWLHQQTTDENLVPFKLTVPWQADMSREQYGEKFRQIQAYLRSGDCYQINLAQR
                     FSAEYQGDEWQAFLSLSRSNQAPFSAFIRLPDNAILSVSPERFLWLENHQVQTRPIKG
                     TLPRLADPEQDRQQAERLANSAKDQAENLMIVDLLRNDIGRVARPGSVRVPELFVVEP
                     FPAVHHLVSTITAILPPECSPTELLRACFPGGSITGAPKVRAMEIIEQLEPYRRNAYC
                     GNIGYISCCGTMDTNITIRTLMTENGKIYCSAGGGIVADSQEQAEYQETFDKVARILP
                     QLGECVIS"
     gene            4875. .5474
                     /gene="y2536"
     CDS             4875. .5474
                     /gene="y2536"
                     /note="residues 10 to 199 of 199 are 55.78 pct identical
                     to residues 8 to 190 of 192 from E. coli K12 : B1813;
                     residues 10 to 199 of 199 are 55.26 pct identical to
                     residues 8 to 190 of 192 from GenPept : >gb|AAL20740.1|
                     (AE008781) putative NTP pyrophosphohydrolase [Salmonella
                     typhimurium LT2]"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAM86092.1"
                     /db_xref="GI:21959407"
                     /translation="MSELITGQYLSEFINRFQLQLPQPDNVLTHSHYFSATNRRAAVL
                     IPIICRPEPTLLLTRRADHLRKHAGQVAFPGGKADPDDQSLISTALREAEEEVAIPAS
                     VVHVLGKLAPLNSSSGYHVTPIVGLVPANIPFYGNDEEVAGLFEIPLHEALSLSRYHS
                     LDIHREGINHRVYLSWYENQFIWGLTATIIRHLAQQVSI"
     gene            6182. .7546
                     /gene="sdaA"
                     /note="y2537"
     CDS             6182. .7546
                     /gene="sdaA"
                     /function="enzyme; degradation of small molecules; amino
                     acids"
                     /note="residues 1 to 453 of 454 are 84.32 pct identical to
                     residues 1 to 453 of 454 from E. coli K12 : B1814"
                     /codon_start=1
                     /transl_table=11
                     /product="L-serine deaminase"
                     /protein_id="AAM86093.1"
                     /db_xref="GI:21959408"
                     /translation="MISVFDMFKIGIGPSSSHTVGPMKAGKQFVDLLITEGLMPSITR
                     VAVDVYGSLSLTGKGHHTDIAIIMGLAGNLPDTVDIDSIPAFIRDVELRQKLMLANGL
                     HEVDFPREGGMVFRSDNLPLHENGMQIHAFAGDEKVLSKTYYSIGGGFIVDEENFGKA
                     SVNDVSVPYPFNSAAEILANCEQTGLSISGMVMQNELAMHSKEEIESYFTAIWQTMRA
                     CIDRGLNTEGVLPGPLRVPRRASALRRLLVSSDKLSSDPMIVIDWVNMFALAVNEENA
                     AGGRVVTAPTNGACGIVPAVLAYYDHFIEPVTPEIFIRYFLASGAIGILYKMNASISG
                     AEVGCQGEVGVACSMAAAGLAELLGASPIQVCIAAEIGMEHNLGLTCDPVAGQVQVPC
                     IERNAIASVKAINSARMAMRRTSEPRVSLDKVIETMFETGKDMNAKYRETSRGGLAIK
                     VQCT"
     gene            complement(7826. .8323)
                     /gene="y2538"
     CDS             complement(7826. .8323)
                     /gene="y2538"
                     /note="residues 15 to 164 of 165 are 47.33 pct identical
                     to residues 1 to 148 of 152 from E. coli K12 : B1007;
                     residues 1 to 165 of 165 are 73.93 pct identical to
                     residues 1 to 165 of 172 from GenPept : >gb|AAC08740.1|
                     (AF021839) 4-hydroxyphenylacetic acid hydroxylase putative
                     coupling protein [Photorhabdus luminescens]"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAM86094.1"
                     /db_xref="GI:21959409"
                     /translation="MSTENQHHLCFRDAMACLSAAVNIVTTDGPAGRCGITATAVCSV
                     TDTPPTLIICINRNSAMNPVFQQNGCLCINVLNHQQEPMARHFAGMTGSSMEERFSWD
                     IWNVGLLGQPLLRDTLASLEGVIEQVQTIGTHLVYLVHITQIVLREQGHGLIYFKRHF
                     HPVMV"

  Query Match             1.7%;  Score 36;  DB 1;  Length 14350;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-08;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1333 TTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 1368
              ||||||||||||||||||||||||||||||||||||
Db       1004 TTGGATGAAGCTGATCGCATGTTGGATATGGGTTTT 969

Search completed: June 10, 2003, 14:05:22
Job time : 5529 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:         June 10, 2003, 16:06:30 ; Search time 87 Seconds
                                          (without alignments)
                                          10288.146 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  3954
Sequence:       1 atgggggatgaagattggga.........tagatgatgagtcatgggat 2172

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:      1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111758_26020/app_query.fasta_1.2311
-DB=SPTREMBL_21 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_119_@runat_05062003_111758_26020 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      SPTREMBL_21:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2001.5	50.6	715	13	Q8QGG8	Q8qgg8 brachydanio
2	1997	50.5	700	13	O42378	O42378 brachydanio
3	1997	50.5	716	13	O42107	O42107 brachydanio
4	1995.5	50.5	715	13	Q8QFU0	Q8qfu0 brachydanio
5	1959.5	49.6	700	13	Q91372	Q91372 xenopus lae
6	1899.5	48.0	617	13	Q90ZF6	Q90zf6 oryzias lat
7	1885	47.7	647	13	Q9PT10	Q9pt10 oncorhynchu
8	1873	47.4	621	13	Q8QHL6	Q8qhl6 oreochromis
9	1865	47.2	645	13	Q9DEG3	Q9deg3 oreochromis
10	1763.5	44.6	662	13	Q9DGR8	Q9dgr8 gallus gall
11	1570	39.7	770	5	Q9GNP1	Q9gnp1 ciona savig
12	1563	39.5	688	5	Q9GNP2	Q9gnp2 ciona savig
13	1522	38.5	659	5	O96069	O96069 ciona intes
14	1522	38.5	669	5	O96068	O96068 ciona intes
15	1497	37.9	394	13	Q8QGD0	Q8qgd0 pantodon bu
16	1489.5	37.7	399	13	Q9I8L8	Q9i8l8 danio dangi
17	1473.5	37.3	397	13	Q8QGD2	Q8qgd2 hyphessobry
18	1460	36.9	398	13	Q8QGD3	Q8qgd3 cyprinus ca
19	1454	36.8	396	13	Q8QGC8	Q8qgc8 oncorhynchu
20	1447	36.6	400	13	Q8QGC9	Q8qgc9 melanotaeni
21	1425.5	36.1	546	5	Q9GV10	Q9gv10 ephydatia f
22	1412	35.7	396	13	Q8QGD1	Q8qgd1 oryzias lat
23	1388	35.1	797	5	Q9GV13	Q9gv13 hydra magni
24	1277	32.3	601	5	O01378	O01378 bombyx mori
25	1257.5	31.8	890	5	Q9GV12	Q9gv12 hydra magni
26	1242.5	31.4	798	5	Q9VHP0	Q9vhp0 drosophila
27	1239.5	31.3	798	5	Q8SXI8	Q8sxi8 drosophila
28	1238	31.3	688	13	O42375	O42375 brachydanio
29	1202	30.4	658	11	Q9QWS9	Q9qws9 mus musculu
30	1179.5	29.8	628	5	Q9GV14	Q9gv14 hydra magni
31	1176	29.7	688	3	Q9P6U9	Q9p6u9 neurospora
32	1165	29.5	617	3	Q8TFK8	Q8tfk8 candida gla
33	1147.5	29.0	726	5	O97031	O97031 dugesia jap
34	1128.5	28.5	641	5	Q9N3F4	Q9n3f4 caenorhabdi
35	1127	28.5	644	5	Q95XY1	Q95xy1 caenorhabdi
36	1100	27.8	633	10	Q9SIN6	Q9sin6 arabidopsis
37	1095.5	27.7	646	10	Q9M2F9	Q9m2f9 arabidopsis
38	1094.5	27.7	763	5	Q22873	Q22873 caenorhabdi
39	1093	27.6	781	5	O97032	O97032 dugesia jap
40	1075	27.2	974	5	Q966L9	Q966l9 caenorhabdi
41	1075	27.2	974	5	Q27376	Q27376 caenorhabdi
42	1069.5	27.0	603	10	Q9M2G5	Q9m2g5 arabidopsis
43	1066.5	27.0	491	5	Q9GV11	Q9gv11 ephydatia f
44	1017.5	25.7	573	5	Q9GV07	Q9gv07 dugesia dor
45	999	25.3	376	5	Q8SXU8	Q8sxu8 drosophila

ALIGNMENTS

```
RESULT 1
Q8QGG8
ID   Q8QGG8      PRELIMINARY;      PRT;   715 AA.
AC   Q8QGG8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa-like protein.
GN   VASA.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB;
RA   Bartfai R., Orban L.;
RT   "Characterization of the genomic locus encoding vasa protein in
RT   zebrafish.";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF461759; AAL89410.1; -.
SQ   SEQUENCE   715 AA;  76795 MW;  602331B3700203AA CRC64;

Alignment Scores:
```

```
Pred. No.:                 1.05e-147      Length:          715
Score:                     2001.50        Matches:         427
Percent Similarity:        66.97%         Conservative:    86
Best Local Similarity:     55.74%         Mismatches:      156
Query Match:               50.62%         Indels:          97
DB:                        13             Gaps:            19

US-09-714-865-15 (1-2172) x Q8QGG8 (1-715)

Qy         10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAG 69
               :::|||||||||   :::    :::|||   :::|||
Db          2 AspAspTrpGluGluAspGlnSerProValValSer------------------------ 13

Qy         70 GATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAA 129
                                                                  |||||
Db         14 ------------------------------------------------CysSerSerGly 17

Qy        130 ATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGG 189
                 |||    :::      |||         ||||||     :::        |||
Db         18 PheGlyLeuGlySerAsnGlySerAspGlyGlyPheLysSerPheTyrThrGlyGlyAla 37

Qy        190 AATTTTGGAAACAGAGATGCT------------------GGTGAGTGTAATAAGCGAGAT 231
                    ||||||      :::                   |||      |||
Db         38 ------GlyAsnAspLysSerAsnSerGluGlyThrGluGlySerSerTrpLysMetThr 55

Qy        232 AATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGA---GGTTTT 288
                 |||     |||     |||    |||    |||    |||    |||    ||||||
Db         56 GlyAspSerPheArgGlyArgGlyGlyArgGlyGlySerArgGlyGlyArgGlyGlyPhe 75

Qy        289 TCA------AACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGT 342
              |||          |||       ::::::        |||    |||    |||
Db         76 SerGlyPheLysSerGluIleAspGluAsnGlySerAspGlyGlyTrp------------ 91

Qy        343 AATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
              |||       |||                               ||||||      |||||||||:::|||
Db         92 AsnGlyGlyGluSerArgGlyLysGlyArgGlyGlyPhe------ArgGlyGlyPheArg 109

Qy        403 GATGGAAATAATTCAGAA-----------------------------GCTTCAGGGCCA 432
                 |||:::      |||                                    |||
Db        110 SerGlySerArgAspGluAsnAspGluAsnGlyAsnAspAspGlyTrpLysGlyGlyGlu 129

Qy        433 TACAGAAGAGGTGGAAGAGGTAGTTTC---CGAGGTTGCCGTGGAGGATTT--------- 480
                 |||      |||||||||    |||       |||    |||||||||||
Db        130 SerArgGlyArgGlyArgGlyGlyPheGlyGlyGlyPheArgGlyGlyPheArgAspGly 149

Qy        481 -----------------------GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGAC 516
                                     |||    |||      ||||||:::         |||
Db        150 GlyAsnGluAspThrGlyArgArgGlyPheGlyArgGluAsnAsnGluAsnGlyAsnAsp 169

Qy        517 GAATGTATGCAGCGC---------ACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTA 567
              |||          |||          ||||||      |||    |||
Db        170 GluGlyGluGlyArgGlyArgGlyArgGlyGlyPheArgGlyGlyPheArg--------- 186

Qy        568 AGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGT 627
                    |||   ||||||    |||      |||             |||      ||||||
Db        187 ------AspGlyGlyGlyAspGluSerGlyLysArg---------GlyPheGlyArgGly 201

Qy        628 GGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCA 687
              |||:::::::|||   |||||||||||      :::              :::   |||
Db        202 GlyPheArgGlyArgAsnGluGluValPheSer---LysValThrThrAlaAspLysLeu 220

Qy        688 GAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCT 747
              :::   ||||||   |||::::::              |||||||||||   |||:::||||||
Db        221 AspGlnGluGlySerGluAsnAla---------GlyProLysValValTyrValProPro 237

Qy        748 CCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGAC 807
              ||||||||||||:::|||   ||||||||:::||||||   |||||||||||||||||||
Db        238 ProProProGluGluGluSerSerIlePheSerHisTyrAlaThrGlyIleAsnPheAsp 257

Qy        808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
              |||||||||   |||||||||:::|||||||||   :::      |||   |||||||::::|||
Db        258 LysTyrAspAspIleLeuValAspValSerGlySerAsnProProLysAlaIleMetThr 277

Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
              ||||||||||||   ||||||    :::|||:::   |||::::::|||:::||||||
Db        278 PheGluGluAlaGlyLeuCysAspSerLeuSerLysAsnValSerLysSerGlyTyrVal 297

Qy        928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
              |||   ||||||||||||:::   |||||||||   ||||||||||||||||||
Db        298 LysProThrProValGlnLysHisGlyIleProIleIleSerAlaGlyArgAspLeuMet 317

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnArg 337

Qy       1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
                 |||   ||||||:::   ||||||:::|||   |||:::|||||||||||   |||
Db        338 PheMetThrAspGlyValAlaAlaSerLysPheSerGluIleGlnGluProGluAlaIle 357

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTGGAAGCCAGAAAATTTTCT 1167
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||:::
Db        358 IleValAlaProThrArgGluLeuIleAsnGlnIleTyrLeuGluAlaArgLysPheAla 377

Qy       1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
              :::||||||||||||||||   ||||||||||||         |||:::::|||
Db        378 TyrGlyThrCysValArgProValValValTyrGlyGlyIleAsnThrGlyTyrThrIle 397

Qy       1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
              |||::::::::::::||||||||:::|||||||||||||||||||||   |||:::
Db        398 ArgGluValLeuLysGlyCysAsnValLeuCysAlaThrProGlyArgLeuHisAspLeu 417

Qy       1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
              ||||||:::   ||||||||||||   :::::::::|||||||||||||||||||||
Db        418 IleGlyArgGlyLysIleGlyLeuSerLysValArgTyrLeuValLeuAspGluAlaAsp 437

Qy       1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
              |||||||||||||||||||   |||||||||:::||||||::::::   |||||||||
Db        438 ArgMetLeuAspMetGlyPheGluProGluMetArgLysLeuValAlaSerProGlyMet 457

Qy       1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
              ||||||||||||:::|||||||||||||||||||||||||||||||:::|||||||||
Db        458 ProSerLysGluGluArgGlnThrLeuMetPheSerAlaThrTyrProGluAspIleGln 477

Qy       1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGT 1527
              |||:::||||||:::|||||||||   :::|||:::|||:::|||||||||   ||||||
Db        478 ArgMetAlaAlaAspPheLeuLysValAspTyrIlePheLeuAlaValGlyValValGly 497

Qy       1528 GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAA 1587
              |||||||||   ||||||:::||||||::::::||||||   |||:::|||||||||:::
Db        498 GlyAlaCysSerAspValGluGlnThrIleValGlnValAspGlnTyrSerLysArgAsp 517

Qy       1588 AAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACT 1647
              :::|||:::|||:::||||||      |||:::||||||||||||||||||||||||
Db        518 GlnLeuLeuGluLeuLeuArgAlaThrGlyAsnGluArgThrMetValPheValGluThr 537

Qy       1648 AAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGT 1707
              |||:::      |||||||||   |||||||||||||||||||||||||||||||||
Db        538 LysArgSerAlaAspPheIleAlaThrPheLeuCysGlnGluLysIleSerThrThrSer 557

Qy       1708 ATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGA 1767
              |||||||||||||||||||||||||||||:::||||||      |||||||||   |||
Db        558 IleHisGlyAspArgGluGlnArgGluArgGluLysAlaLeuSerAspPheArgLeuGly 577

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              :::||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        578 GlnCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluGlnVal 597

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              |||||||||:::|||||||||:::||||||:::||||||||||||||||||||||||
Db        598 GlnHisValValAsnPheAspMetProSerSerIleAspGluTyrValHisArgIleGly 617

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              |||||||||||||||||||||||||||||||:::|||||||||:::      |||||||||
Db        618 ArgThrGlyArgCysGlyAsnThrGlyArgAlaValSerPhePheAsnProGluSerAsp 637

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
```

```
                   ||||||:::  ||||||||||||||:::  |||||||||   ||||||
Db        638 ThrProLeuAlaArgSerLeuValLysValLeuSerGlyAlaGlnGlnValValProLys 657

Qy       2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGT---AGTACAAGA 2064
              ||||||||||:::||||||||   |||::::||||||||   :::      |||   :::|||   :::   |||
Db        658 TrpLeuGluGluValAlaPheSerAlaHis------GlyThrThrGlyPheAsnProArg 675

Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGG 2124
              |||   |||||||||||||   |||:::||||||||||   |||
Db        676 GlyLysValPheAlaSerThrAspSerArgLysGlyGlySer----------------- 689

Qy       2125 TTTCTTCTTCACGAGCTCCC----------------------------AATCCAGTA 2154
              |||   |||         |||
Db        690 PheLysSerAspGluProProProSerGlnThrSerAlaProSerAlaAlaAlaAlaAla 709

Qy       2155 GATGATGAGTCATGGGAT 2172
              |||||||||   |||:::
Db        710 AspAspGluGluTrpGlu 715
```

RESULT 2
O42378

```
ID   O42378      PRELIMINARY;      PRT;   700 AA.
AC   O42378;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   RNA helicase (DEAD box).
GN   VASA OR VLG.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98025484; PubMed=9376327;
RA   Olsen L.C., Aasland R., Fjose A.;
RT   "A vasa-like gene in Zebrafish identifies putative primordial germ
RT   cells.";
RL   Mech. Dev. 66:95-105(1997).
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; Y12007; CAA72735.1; -.
DR   HSSP; Q58083; 1HV8.
DR   ZFIN; ZDB-GENE-990415-272; vasa.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   700 AA;  75312 MW;  726B56A90D6C2DB3 CRC64;
```

Alignment Scores:

Pred. No.:	2.34e-147	Length:	700
Score:	1997.00	Matches:	423
Percent Similarity:	67.37%	Conservative:	89
Best Local Similarity:	55.66%	Mismatches:	148
Query Match:	50.51%	Indels:	100
DB:	13	Gaps:	20

US-09-714-865-15 (1-2172) x O42378 (1-700)

```
Qy         10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAG 69
              :::|||||||||   :::   :::|||   :::|||
Db          2 AspAspTrpGluGluAspGlnSerProValValSerCys-----------------Ser 15

Qy         70 GATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAA 129
                 :::   |||         |||                  |||:::|||||||
Db         16 SerGlyPheGlyGlyAlaGlyAsnAspLys-----------------SerAsnSerGlu 29

Qy        130 ATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGG 189
               :::|||   |||   :::                      :::|||
Db         30 GlyThrGluGlySerSerTrpLys----------------------MetThrGlyAsp 41

Qy        190 AATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGT 249
              :::|||   |||                                          |||
Db         42 SerPheArgGlyArg---------------------------------------Gly 47

Qy        250 GGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCA------AACAGCAGGTTT 303
              |||   |||   |||         |||      ||||||||           |||
Db         48 GlyArgGly---GlySerArgGlyGlyArgGlyGlyPheSerGlyPheLysSerGluIle 66

Qy        304 GAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCA 363
              ::::::      |||   |||   |||            |||      |||
Db         67 AspGluAsnGlySerAspGlyGlyTrp-----------AsnGlyGlyGluSerArgGly 82

Qy        364 ACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA--- 420
                       ||||||      |||||||||:::|||   |||:::      |||
Db         83 ArgGlyArgGlyGlyPhe------ArgGlyGlyPheArgSerGlySerArgAspGluAsn 100

Qy        421 -------------------------GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGT 453
                                       |||      |||      |||||||||
Db        101 AspGluAsnGlyAsnAspAspGlyTrpLysGlyGlyGluSerArgGlyArgGlyArgGly 120

Qy        454 AGTTTC---CGAGGTTGCCGTGGAGGATTT---------------------------- 480
                 |||      |||   ||||||||||||
Db        121 GlyPheGlyGlyGlyPheArgGlyGlyPheArgAspGlyGlyAsnGluAspThrGlyArg 140

Qy        481 ---GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG--------- 528
                 |||   |||      ||||||:::         ||||||      :::
Db        141 ArgGlyPheGlyArgGluAsnAsnGluAsnGlyAsnAspGluGlyGlyGluGlyArgGly 160

Qy        529 ---CGCACTGGTGGCCTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGT 585
                 ||||||      |||      |||                  |||   |||
Db        161 ArgGlyArgGlyGlyPheArgGlyGlyPheArg--------------AspGlyGlyGly 175

Qy        586 GATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAAT 645
              |||   |||      |||         |||      |||||||||::::::|||   |||
Db        176 AspGluSerGlyLysArg---------GlyPheGlyArgGlyGlyPheArgGlyArgAsn 192

Qy        646 GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
              |||||||||   :::                 :::   |||   :::   ||||||   |||
Db        193 GluGluValPheSer---LysValThrThrAlaAspLysLeuAspGlnGluGlySerGlu 211

Qy        706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAG 765
              ::::::         ||||||||||||   |||:::||||||||||||||||||:::|||
Db        212 AsnAla---------GlyProLysValValTyrValProProProProProGluGluGlu 228

Qy        766 GACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTT 825
                 |||||||||:::||||||   |||||||||||||||||||||||||   ||||||
Db        229 SerSerIlePheSerHisTyrAlaThrGlyIleAsnPheAspLysTyrAspAspIleLeu 248

Qy        826 GTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTC 885
              |||:::|||||||||   :::      |||   |||||||:::|||||||||||||   |||
Db        249 ValAspValSerGlySerAsnProProLysAlaIleMetThrPheGluGluAlaGlyLeu 268

Qy        886 TGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAA 945
              |||   :::|||:::   |||::::::|||:::||||||   |||   ||||||||||||
Db        269 CysAspSerLeuSerLysAsnValSerLysSerGlyTyrValLysProThrProValGln 288

Qy        946 AAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGG 1005
              |||:::   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        289 LysHisGlyIleProIleIleSerAlaGlyArgAspLeuMetAlaCysAlaGlnThrGly 308

Qy       1006 TCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATA 1065
              |||||||||||||||||||||||||||||||||||         |||   ||||||:::
Db        309 SerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnArgPheMetThrAspGlyVal 328

Qy       1066 ACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGA 1125
                 ||||||:::|||   |||:::||||||||||||   ||||||||||||||||||
Db        329 AlaAlaSerLysPheSerGluIleGlnGluProGluAlaIleIleValAlaProThrArg 348

Qy       1126 GAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGA 1185
```

```
                 |||||:::||||||||||||||||||||||||||||||:::::::|||||||||||||
Db          349 GluLeuIleAsnGlnIleTyrLeuGluAlaArgLysPheAlaTyrGlyThrCysValArg 368

Qy         1186 GCTGTTGTTATATATGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGC 1245
                    ||||||:::||||||||           |||:::::::||||||:::::::::::::|||
Db          369 ProValValValTyrGlyGlyIleAsnThrGlyTyrThrIleArgGluValLeuLysGly 388

Qy         1246 TGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATT 1305
                |||||:::|||||||||||||||||||||||    |||:::||||||:::    ||||||
Db          389 CysAsnValLeuCysAlaThrProGlyArgLeuHisAspLeuIleGlyArgGlyLysIle 408

Qy         1306 GGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGT 1365
                ||||||    ::::::::||||||||||||||||||||||||||||||||||||||||||
Db          409 GlyLeuSerLysValArgTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGly 428

Qy         1366 TTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGC 1425
                |||    |||||||||:::||||||::::::    ||||||||||||||||||||:::|||
Db          429 PheGluProGluMetArgLysLeuValAlaSerProGlyMetProSerLysGluLysArg 448

Qy         1426 CAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTT 1485
                ||||||||||||||||||||||||:::||||||:::|||||||||:::||||||:::|||
Db          449 GlnThrLeuMetPheSerAlaThrTyrProGluAspIleGlnArgMetAlaAlaAspPhe 468

Qy         1486 TTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTT 1545
                ||||||    :::|||:::|||:::|||||||||    |||||||||||||||    ||||||
Db          469 LeuLysValAspTyrIlePheLeuAlaValGlyValValGlyGlyAlaCysSerAspVal 488

Qy         1546 CAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTG 1605
                :::|||||:::::::||||||    |||:::||||||||||:::::::|||:::|||:::|||
Db          489 GluGlnThrIleValGlnValAspGlnTyrSerLysArgAspGlnLeuLeuGluLeuLeu 508

Qy         1606 CGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTT 1665
                |||          |||:::||||||||||||||||||||||||||||||:::    ||||||||
Db          509 ArgAlaThrGlyAsnGluArgThrMetValPheValGluThrLysArgSerAlaAspPhe 528

Qy         1666 ACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAA 1725
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          529 IleAlaThrPheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGlu 548

Qy         1726 CAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTT 1785
                ||||||||||||:::||||||    |||||||||    |||    ||||||||||||||
Db          549 GlnArgGluArgGluLysAlaLeuSerAspPheArgLeuGlyHisCysProValLeuVal 568

Qy         1786 GCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTT 1845
                ||||||||||||||||||||||||||||||||||||    ||||||||||||:::||||||
Db          569 AlaThrSerValAlaAlaArgGlyLeuAspIleGluGlnValGlnHisValValAsnPhe 588

Qy         1846 GATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGG 1905
                |||:::||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db          589 AspMetProSerSerIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGly 608

Qy         1906 AATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCT 1965
                |||||||||||||||:::|||||||||:::    |||||||||         ||||||:::
Db          609 AsnThrGlyArgAlaValSerPhePheAsnProGluSerAspThrProLeuAlaArgSer 628

Qy         1966 CTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCC 2025
                |||||||||||||:::    |||||||||    ||||||    ||||||||||||:::|||
Db          629 LeuValLysValLeuSerGlyAlaGlnGlnValValProLysTrpLeuGluGluValAla 648

Qy         2026 TTTAGTACATACATTCCTGGCTTCAGTGGT---AGTACAAGAGGAAACGTGTTTGCATCA 2082
                ||||||       :::       |||    :::|||     :::    ||||||    ||||||||||
Db          649 PheSerAlaHis------GlyThrThrGlyPheAsnProArgGlyLysValPheAlaSer 666

Qy         2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                    |||:::|||||||||    |||                  |||    |||
Db          667 ThrAspSerArgLysGlyGlySer------------------PheLysSerAspGluPro 680

Qy         2143 CCC------------------------------AATCCAGTAGATGATGAGTCATGGGAT 2172
                |||                                   |||||||||    |||:::
Db          681 ProProSerGlnThrSerAlaProSerAlaAlaAlaAlaAlaAspAspGluGluTrpGlu 700
```

RESULT 3
O42107

```
ID   O42107      PRELIMINARY;      PRT;   716 AA.
AC   O42107;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Vas.
GN   VASA.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97417583; PubMed=9272956;
RA   Yoon C., Kawakami K., Hopkins N.;
RT   "Zebrafish vasa homologue RNA is localized to the cleavage planes of
RT   2-and 4-cell-stage embryos and is expressed in the primordial germ
RT   cells.";
RL   Development 124:3157-3166(1997).
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AB005147; BAA22535.1; -.
DR   HSSP; Q58083; 1HV8.
DR   ZFIN; ZDB-GENE-990415-272; vasa.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   716 AA;  77004 MW;  BB74923B95C7AC5A CRC64;

Alignment Scores:
Pred. No.:               2.35e-147     Length:         716
Score:                   1997.00       Matches:        427
Percent Similarity:      66.75%        Conservative:   85
Best Local Similarity:   55.67%        Mismatches:     157
Query Match:             50.51%        Indels:         98
DB:                      13            Gaps:           19

US-09-714-865-15 (1-2172) x O42107 (1-716)

Qy           10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAGAAG 69
                :::|||||||||    :::    :::|||    :::|||
Db            2 AspAspTrpGluGluAspGlnSerProValValSer---------------------- 13

Qy           70 GATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAA 129
                                                              ||||||
Db           14 ---------------------------------------------CysSerSerGly 17

Qy          130 ATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGG 189
                         |||    :::      |||         ||||||    :::      |||
Db           18 PheGlyLeuGlySerAsnGlySerAspGlyGlyPheLysSerPheTyrThrGlyGlyAla 37

Qy          190 AATTTTGGAAACAGAGATGCT-----------------GGTGAGTGTAATAAGCGAGAT 231
                      ||||||    :::                   |||         |||
Db           38 ------GlyAsnAspLysSerAsnSerGluGlyThrGluGlySerSerTrpLysMetThr 55

Qy          232 AATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGA---GGTTTT 288
                   |||      |||      |||   |||   .|||   |||   |||   ||||||
Db           56 GlyAspSerPheArgGlyArgGlyGlyArgGlyGlySerArgGlyGlyArgGlyGlyPhe 75

Qy          289 TCA------AACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGT 342
                |||         |||      ::::::      |||   |||   |||
Db           76 SerGlyPheLysSerGluIleAspGluAsnGlySerAspGlyGlyTrp----------- 91

Qy          343 AATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
                   |||      |||                  ||||||         |||||||||:::|||
Db           92 AsnGlyGlyGluSerArgGlyArgGlyArgGlyGlyPhe------ArgGlyGlyPheArg 109
```

```
Qy        403 GATGGAAATAATTCAGAA----------------------------GCTTCAGGGCCA 432
Db        110 SerGlySerArgAspGluAsnAspGluAsnArgAsnAspAspGlyTrpLysGlyGlyGlu 129

Qy        433 TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGC---CGTGGAGGATTT--------- 480
Db        130 SerArgGlyArgGlyArgGlyGlyPheGlyGlySerPheArgGlyGlyPheArgAspGly 149

Qy        481 ------------------------GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGAC 516
Db        150 GlyAsnGluAspThrGlyArgArgGlyPheGlyArgGluAsnAsnGluAsnGlyAsnAsp 169

Qy        517 GAATGTATGCAG------------CGCACTGGTGGCCTTTTGGTTCTAGAAGACCAGTA 564
Db        170 GluGlyGlyGluGlyArgGlyArgGlyArgGlyGlyPheArgGlyGlyPheArg------ 187

Qy        565 TTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGA 624
Db        188 ---------AspGlyGlyGlyAspGluSerGlyLysArg---------GlyPheGlyArg 201

Qy        625 GGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAG 684
Db        202 GlyGlyPheArgGlyArgAsnGluGluValPheSer---LysValThrThrAlaAspLys 220

Qy        685 TCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCC 744
Db        221 LeuAspGlnGluGlySerGluAsnAla---------GlyProLysValValTyrValPro 237

Qy        745 CCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTC 804
Db        238 ProProProProGluGluGluSerSerIlePheSerHisTyrAlaThrGlyIleAsnPhe 257

Qy        805 GACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTG 864
Db        258 AspLysTyrAspAspIleLeuValAspValSerGlySerAsnProProLysAlaIleMet 277

Qy        865 ACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTAT 924
Db        278 ThrPheGluGluAlaGlyLeuCysAspSerLeuSerLysAsnValSerLysSerGlyTyr 297

Qy        925 ACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTG 984
Db        298 ValLysProThrProValGlnLysHisGlyIleProIleIleSerAlaGlyArgAspLeu 317

Qy        985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCT 1044
Db        318 MetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuGln 337

Qy       1045 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
Db        338 ArgPheMetThrAspGlyValAlaAlaSerLysPheSerGluIleGlnGluProGluAla 357

Qy       1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164
Db        358 IleIleValAlaProThrArgGluLeuIleAsnGlnIleTyrLeuGluAlaArgLysPhe 377

Qy       1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA 1224
Db        378 AlaTyrGlyThrCysValArgProValValValTyrGlyGlyIleAsnThrGlyTyrThr 397

Qy       1225 ATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGAT 1284
Db        398 IleArgGluValLeuLysGlyCysAsnValLeuCysAlaThrProGlyArgLeuHisAsp 417

Qy       1285 ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCT 1344
Db        418 LeuIleGlyArgGlyLysIleGlyLeuSerLysValArgTyrLeuValLeuAspGluAla 437

Qy       1345 GATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
Db        438 AspArgMetLeuAspMetGlyPheGluProGluMetArgLysLeuValAlaSerProGly 457

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
Db        458 MetProSerLysGluGluArgGlnThrLeuMetPheSerAlaThrTyrProGluAspIle 477

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
Db        478 GlnArgMetAlaAlaAspPheLeuLysValAspTyrIlePheLeuAlaValGlyValVal 497

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
Db        498 GlyGlyAlaCysSerAspValGluGlnThrValValGlnValAspGlnTyrSerLysArg 517

Qy       1585 GAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAA 1644
Db        518 AspGlnLeuLeuGluLeuLeuArgAlaThrGlyAsnGluArgThrMetValPheValGlu 537

Qy       1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACA 1704
Db        538 ThrLysArgSerAlaAspPheIleAlaThrPheLeuCysGlnGluLysIleSerThrThr 557

Qy       1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
Db        558 SerIleHisGlyAspArgGluGlnArgGluArgGluLysAlaLeuSerAspPheArgLeu 577

Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824
Db        578 GlyHisCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluGln 597

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884
Db        598 ValGlnHisValValAsnPheAspMetProSerSerIleAspGluTyrValHisArgIle 617

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCG 1944
Db        618 GlyArgThrGlyArgCysGlyAsnThrGlyArgAlaValSerPhePheAsnProGluSer 637

Qy       1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
Db        638 AspThrProLeuAlaArgSerLeuValLysValLeuSerGlyAlaGlnGlnValValPro 657

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGT---AGTACA 2061
Db        658 LysTrpLeuGluGluValAlaPheSerAlaHis------GlyThrThrGlyPheAsnPro 675

Qy       2062 AGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCT 2121
Db        676 ArgGlyLysValPheAlaSerThrAspSerArgLysGlyGlySer--------------- 690

Qy       2122 GGGTTTTCTTCTTCACGAGCTCCC-----------------------------AATCCA 2151
Db        691 ---PheLysSerAspGluProProProSerGlnThrSerAlaProSerAlaAlaAlaAla 709

Qy       2152 GTAGATGATGAGTCATGGGAT 2172
Db        710 AlaAspAspGluGluTrpGlu 716
```

```
RESULT 4
Q8QFU0
ID   Q8QFU0      PRELIMINARY;      PRT;   715 AA.
AC   Q8QFU0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa-like protein.
GN   VLG.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Vatland A., Olsen L.C.;
RT   "Establishment of transgenic zebrafish lines expressing green
```

```
RT   fluorescent protein in their germ cells.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ311625; CAC84069.1; -.
SQ   SEQUENCE   715 AA;  76880 MW;  3CE4E03F02073BA4 CRC64;

Alignment Scores:
Pred. No.:              3.08e-147      Length:          715
Score:                  1995.50        Matches:         426
Percent Similarity:     66.88%         Conservative:    89
Best Local Similarity:  55.32%         Mismatches:      150
Query Match:            50.47%         Indels:          105
DB:                     13             Gaps:            21

US-09-714-865-15 (1-2172) x Q8QFU0 (1-715)

Qy        10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA--------- 60
             :::||||||||  :::   :::|||  :::|||
Db         2 AspAspTrpGluGluAspGlnSerProValValSerCysSerSerGlyPheGlyLeuGly 21

Qy        61 --------------------TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAAT 99
                                 |||   |||   |||:::|||           |||
Db        22 SerAsnGlySerAspGlyGlyPhe---LysSerPheTyrThrGlyAlaGlyAsnAspLys 40

Qy       100 TTTAACAGGACTCCAGCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCAT 159
                              |||:::||||||   :::|||   |||   :::
Db        41 -----------------SerAsnSerGluGlyThrGluGlySerSerTrpLys------ 52

Qy       160 TTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGT 219
                                 :::|||   :::|||      |||
Db        53 -----------------MetThrGlyAspSerPheArgGlyArg-------------- 61

Qy       220 AATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAAC 279
                                       ||||||   |||   |||      |||
Db        62 --------------------------GlyGlyArgGly---GlySerArgGlyGlyArg 71

Qy       280 AGAGGTTTTTCA------AACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGA 333
                |||||||||      |||      ::::::     |||   |||   |||
Db        72 GlyGlyPheSerGlyPheLysSerGluIleAspGluAsnGlySerAspGlyGlyTrp--- 90

Qy       334 GAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGC 393
                      |||   |||                ||||||         ||||||
Db        91 ---------AsnGlyGlyGluSerArgGlyArgGlyArgGlyGlyPhe------ArgGly 105

Qy       394 GGCTATCGAGATGGAAATAATTCAGAA-----------------------------GCT 423
             |||:::|||   |||:::      |||
Db       106 GlyPheArgSerGlySerArgAspGluAsnAspGluAsnGlyAsnAspAspGlyTrpLys 125

Qy       424 TCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGC---CGTGGAGGATTT 480
                |||         |||      |||||||||   |||   |||      |||||||||||
Db       126 GlyGlyGluSerArgGlyArgGlyArgGlyGlyPheGlyGlySerPheArgGlyGlyPhe 145

Qy       481 ---------------------------------GGTCTAGGAAGTCCAAATAATGACTTA 507
                                              |||   |||      ||||||:::
Db       146 ArgAspGlyGlyAsnGluAspThrGlyArgArgGlyPheGlyArgGluAsnAsnGluAsn 165

Qy       508 GACCCAGACGAATGTATGCAG------------CGCACTGGTGGCCTTTTGGTTCTAGA 555
                   ||||||      :::                     ||||||      |||
Db       166 GlyAsnAspGluGlyGlyGluGlyArgGlyArgGlyArgGlyGlyPheArgGlyGlyPhe 185

Qy       556 AGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGA 615
             |||                     |||   ||||||   |||      |||      |||
Db       186 Arg---------------AspGlyGlyGlyAspGluSerGlyLysArg---------Gly 197

Qy       616 AGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAAT 675
                   |||||||||::::::|||   |||||||||||      :::
Db       198 PheGlyArgGlyGlyPheArgGlyArgAsnGluGluValPheSer---LysValThrThr 216

Qy       676 TCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACC 735
             :::   |||   :::   ||||||   |||::::::          |||||||||||
Db       217 AlaAspLysLeuAspGlnGluGlySerGluAsnAla---------GlyProLysValVal 233

Qy       736 TACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGC 795
             |||:::||||||||||||||||||:::|||     |||||||||:::||||||   ||||||
Db       234 TyrValProProProProProGluGluGluSerSerIlePheSerHisTyrAlaThrGly 253

Qy       796 ATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCA 855
             |||||||||||||||||||   |||||||||:::|||||||||   :::   |||
Db       254 IleAsnPheAspLysTyrAspAspIleLeuValAspValSerGlySerAsnProProLys 273

Qy       856 GCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAA 915
             ||||||:::||||||||||||||   ||||||   :::|||:::   |||::::::|||
Db       274 AlaIleMetThrPheGluGluAlaGlyLeuCysAspSerLeuSerLysAsnValSerLys 293

Qy       916 GCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGA 975
             :::||||||   |||   |||||||||||||||:::   ||||||||||||   ||||||
Db       294 SerGlyTyrValLysProThrProValGlnLysHisGlyIleProIleIleSerAlaGly 313

Qy       976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       314 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 333

Qy      1036 ATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA 1095
             ||||||         |||   ||||||:::   ||||||:::|||      |||:::||||||
Db       334 IleLeuGlnArgPheMetThrAspGlyValAlaAlaSerLysPheSerGluMetGlnGlu 353

Qy      1096 CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTGGAAGCC 1155
             ||||||   ||||||||||||||||||||||||||:::|||||||||||||||||||||
Db       354 ProGluAlaIleIleValAlaProThrArgGluLeuIleAsnGlnIleTyrLeuGluAla 373

Qy      1156 AGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTG 1215
             |||||||||::::::|||||||||||||||||   ||||||:::|||||||||
Db       374 ArgLysPheAlaTyrGlyThrCysValArgProValValValTyrGlyGlyIleAsnThr 393

Qy      1216 GGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGA 1275
             |||::::::||||||::::::::::::||||||||:::||||||||||||||||||||||||
Db       394 GlyTyrThrIleArgGluValLeuLysGlyCysAsnValLeuCysAlaThrProGlyArg 413

Qy      1276 CTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTG 1335
             |||   |||:::||||||:::   ||||||||||||   :::::::::||||||||||||
Db       414 LeuHisAspLeuIleGlyArgGlyLysIleGlyLeuSerLysValArgTyrLeuValLeu 433

Qy      1336 GATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT 1395
             ||||||||||||||||||||||||||||||||||   |||||||||:::||||||::::::
Db       434 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGluMetArgLysLeuValAla 453

Qy      1396 TGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCA 1455
                ||||||||||||||||||||:::||||||||||||||||||||||||||||:::|||
Db       454 SerProGlyMetProSerLysGluGluArgGlnThrLeuMetPheSerAlaThrTyrPro 473

Qy      1456 GAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT 1515
             |||:::|||||||||||||||:::|||||||||      :::|||:::|||:::||||||
Db       474 GluAspIleGlnArgMetAlaAlaAspPheLeuLysValAspTyrIlePheLeuAlaVal 493

Qy      1516 GGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTC 1575
             |||   |||||||||||||||   ||||||:::||||||::::::||||||   |||:::
Db       494 GlyValValGlyGlyAlaCysSerAspValGluGlnThrIleValGlnValAspGlnTyr 513

Qy      1576 TCAAAAAGAGAAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTC 1635
             |||||||||::::::|||:::|||:::||||||         |||:::||||||||||||||
Db       514 SerLysArgAspGlnLeuLeuGluLeuLeuArgAlaThrGlyAsnGluArgThrMetVal 533

Qy      1636 TTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATA 1695
             |||||||||||||||:::   |||||||||   ||||||||||||||||||||||||||||
Db       534 PheValGluThrLysArgSerAlaAspPheIleAlaThrPheLeuCysGlnGluLysIle 553

Qy      1696 TCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGAT 1755
             ||||||||||||||||||||||||||||||||||||||||||||:::||||||   |||
Db       554 SerThrThrSerIleHisGlyAspArgGluGlnArgGluArgGluLysAlaLeuSerAsp 573

Qy      1756 TTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGAT 1815
             ||||||   |||   |||||||||||||||||||||||||||||||||||||||||||||
Db       574 PheArgLeuGlyHisCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAsp 593

Qy      1816 ATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTT 1875
             ||||||   ||||||||||||:::|||||||||:::||||||:::||||||||||||||||
Db       594 IleGluGlnValGlnHisValValAsnPheAspMetProSerSerIleAspGluTyrVal 613
```

```
Qy       1876 CATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGAT 1935
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||:::
Db        614 HisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAlaValSerPhePheAsn 633

Qy       1936 CTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAG 1995
                 |||||||||      ||||||:::   ||||||||||||||:::   |||||||||
Db        634 ProGluSerAspThrProLeuAlaArgSerLeuValLysValLeuSerGlyAlaGlnGln 653

Qy       1996 GATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGT 2055
                 ||||||   ||||||||||||:::||||||||||    :::      |||    :::|||
Db        654 ValValProLysTrpLeuGluGluValAlaPheSerAlaHis------GlyThrThrGly 671

Qy       2056 ---AGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTG 2112
                 :::   ||||||   ||||||||||||   |||:::||||||||||   |||
Db        672 PheAsnProArgGlyLysValPheAlaSerThrAspSerArgLysGlyGlySer------ 689

Qy       2113 AACACAGCTGGGTTTTCTTCTTCACGAGCTCCC--------------------------- 2145
                          |||   |||         |||
Db        690 -----------PheLysSerAspGluProProProSerGlnThrSerAlaProSerAla 705

Qy       2146 ---AATCCAGTAGATGATGAGTCATGGGAT 2172
                          |||||||||   |||:::
Db        706 AlaAlaAlaAlaAspAspGluGluTrpGlu 715
```

RESULT 5
Q91372

```
ID   Q91372      PRELIMINARY;      PRT;    700 AA.
AC   Q91372;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   DEAD box protein.
GN   VLG1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94200507; PubMed=8150200;
RA   Komiya T., Itoh K., Ikenishi K., Furusawa M.;
RT   "Isolation and characterization of a novel gene of the DEAD box
RT   protein family which is specifically expressed in germ cells of
RT   Xenopus laevis.";
RL   Dev. Biol. 162:354-363(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Komiya T.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AF046043; AAC03114.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000685; RuBisCO_large.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS00157; RUBISCO_LARGE; UNKNOWN_1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   700 AA;  78240 MW;  C504ECA38EFB0B7E CRC64;

Alignment Scores:
Pred. No.:               1.98e-144     Length:          700
Score:                   1959.50       Matches:         409
Percent Similarity:      68.17%        Conservative:    105
Best Local Similarity:   54.24%        Mismatches:      153
Query Match:             49.56%        Indels:          87
DB:                      13            Gaps:            18

US-09-714-865-15 (1-2172) x Q91372 (1-700)

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              :::|||:::|||:::   ||||||            :::|||||||||   |||
Db          2 GluGluAsnTrpAspThrGluIleGluThrGluLysProThrTyrValProAsnPhe--- 20

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
                       ||||||   ||||||:::      :::      |||::::::
Db         21 ---SerThrLeuGluThrGluAsnThrAspAsnTyr------SerAlaTyrSerAsnAsn 37

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ::::::::::::                                    :::   |||
Db         38 AspIleAsnAsn-----------------------------GlnAsnTyrAspSerGlu 47

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              |||:::||||||||||||      |||      ::::::|||         |||
Db         48 ArgSerPheGlyAsnArg---GlyGlyTyrArgSerGluArgSerArgProSerAsnPhe 66

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
                                             |||||||||         ||||||   |||
Db         67 -----------------------------AsnArgGly---------SerArgThrGlu 73

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------------- 351
                 |||      ||||||            :::::::::
Db         74 ArgGlyArgGlyArgGlyPheGlyThrAsnArgAsnAspAsnTyrSerSerGluArgAsp 93

Qy        352 ---------GAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
                       :::|||      :::   |||||||||      |||||||||||
Db         94 ValPheGlyAspAspGluArgAspGlnArgArgGlyPheProGlyArgGlyGlyTyrAsn 113

Qy        403 GATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGA 462
                    :::   :::            :::|||      |||||||||   ||||||
Db        114 GlyAsnGluAspGlyGlnLysProAsnAlaPheArg------GlyArgGlyGlyPheArg 131

Qy        463 ---------------------------GGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCA 495
                                         |||   ||||||||||||         :::
Db        132 AsnGluAsnGluGlnArgArgGlyPheGlyGluArgGlyGlyPheArgSerGluAsnGly 151

Qy        496 AATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTGGTTCTAGA 555
                    :::   |||                              |||   ||||||:::
Db        152 GlnArgAsnPheAsp---------------------AsnArgGlyAspPheGlyAsnSer 164

Qy        556 ---------------AGACCAGTATTAAGTGGCACAGGT---------AATGGTGATACT 591
                             ||||||      |||   |||         |||   ||||||
Db        165 GlyGluGluGluAspArgProArgSerTyrGlyArgGlyGlyPheAsnAsnSerAspThr 184

Qy        592 TCT-----------CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGT 639
                           :::      |||   ||||||:::   ||||||||||||||
Db        185 GlyGlyArgGlyArgArgGlyGlyArgGlyGlyGlySerGlnTyrGlyGlyTyrLysGly 204

Qy        640 TTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGA 699
                 ||||||||||||            |||         |||   ||||||:::   ||||||
Db        205 ArgAsnGluGluVal------------GlyValGluSerGlyLysSerGlnGluGluGly 220

Qy        700 GGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAG 759
                 |||      |||      :::      ||||||||||||||||||||||||||||:::
Db        221 AsnGlu---LysAspGluLysProLysLysValThrTyrIleProProProProProAsp 239

Qy        760 GATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACT 819
                 ||||||:::||||||      ||||||:::||||||||||||||||||||||||
Db        240 GlyGluAspAsnIlePheArgGlnTyrGlnSerGlyIleAsnPheAspLysTyrAspGlu 259

Qy        820 ATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCT 879
              |||||||||:::|||:::|||   |||   ||||||||||||||||||||||||||||||
Db        260 IleLeuValAspValThrGlyLysAspValProProAlaIleLeuThrPheGluGluAla 279

Qy        880 AATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCT 939
              ||||||||||||:::||||||      |||:::|||:::|||||||||   ||||||||||||
Db        280 AsnLeuCysGluThrLeuArgArgAsnValAlaArgAlaGlyTyrValLysLeuThrPro 299

Qy        940 GTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAA 999
```

```
Db        300 ValGlnLysHisSerIleProIleIleMetAlaGlyArgAspLeuMetAlaCysAlaGln 319

Qy       1000 ACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGAT 1059
Db        320 ThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerTyrMetMetAsnGlu 339

Qy       1060 GGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCA 1119
Db        340 GlyIleThrAlaSerGlnTyrLeuGlnLeuGlnGluProGluAlaIleIleIleAlaPro 359

Qy       1120 ACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGT 1179
Db        360 ThrArgGluLeuIleAsnGlnIleTyrLeuAspAlaArgLysPheSerTyrGlyThrCys 379

Qy       1180 GTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTA 1239
Db        380 ValArgProValValValTyrGlyGlyIleGlnProValHisAlaMetArgAspValGlu 399

Qy       1240 CAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAA 1299
Db        400 LysGlyCysAsnIleLeuCysAlaThrProGlyArgLeuLeuAspIleValSerLysGlu 419

Qy       1300 AAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGAT 1359
Db        420 LysIleGlyLeuSerLysLeuArgTyrLeuValLeuAspGluAlaAspArgMetLeuAsp 439

Qy       1360 ATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAA 1419
Db        440 MetGlyPheAlaProGluIleGluLysLeuMetThrLysProGlyMetProThrLysGlu 459

Qy       1420 CAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCA 1479
Db        460 LysArgGlnThrLeuMetPheSerAlaThrTyrProGluGluIleArgArgLeuAlaSer 479

Qy       1480 GAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGA 1539
Db        480 AsnTyrLeuLysSerGluHisLeuPheValValValGlyLeuValGlyGlyAlaCysSer 499

Qy       1540 GATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAA 1599
Db        500 AspValAlaGlnThrValLeuGluMetArgGluAsnGlyLysMetGluLysLeuLeuGlu 519

Qy       1600 ATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCA 1659
Db        520 IleLeuLysSerSerGluLysGluArgThrMetIlePheValAsnThrLysLysLysAla 539

Qy       1660 GATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGAT 1719
Db        540 AspPheIleAlaGlyTyrLeuCysGlnGluLysPheSerSerThrSerIleHisGlyAsp 559

Qy       1720 CGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTT 1779
Db        560 ArgGluGlnTyrGlnArgGluSerAlaLeuTrpAspPheArgThrGlyLysCysThrVal 579

Qy       1780 CTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATC 1839
Db        580 IleValCysThrAlaValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIle 599

Qy       1840 AATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
Db        600 AsnTyrAspValProLysGluValAspGluTyrValHisArgIleGlyArgThrGlyArg 619

Qy       1900 TGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCA 1959
Db        620 CysGlyAsnThrGlyLysAlaThrSerPhePheAsnValGlnAspAspHisValIleAla 639

Qy       1960 CAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAA 2019
Db        640 ArgProLeuValLysIleLeuThrAspAlaHisGlnGluValProAlaTrpLeuGluGlu 659

Qy       2020 ATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCA 2079
Db        660 IleAla---------------------PheGlyGlyHisGlyAlaLeuAsnSerPheTyr 672

Qy       2080 TCAGTTGATACC---------AGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCT 2130
Db        673 AlaAlaAspSerMetGlyGluGlnAlaGlyGlyAsnAlaValThrThrProSerPheAla 692

Qy       2131 TCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
Db        693 GlnGlu------------------GluGluAlaSerTrpAsp 700
```

```
RESULT 6
Q90ZF6
ID   Q90ZF6      PRELIMINARY;      PRT;   617 AA.
AC   Q90ZF6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   VASA.
GN   OLVAS.
OS   Oryzias latipes (Medaka fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shinomiya A., Tanaka M., Kobayashi T., Nagahama Y., Hamaguchi S.;
RT   "The vasa-like gene, olvas, identifies the migration path of
RT   primordial germ cells during embryonic body formation stage in the
RT   medaka, Oryzias latipes.";
RL   Development 42:317-326(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21126920; PubMed=11226275;
RA   Tanaka M., Kinoshita M., Kobayashi D., Nagahama Y.;
RT   "Establishment of medaka (Oryzias latipes) transgenic lines with the
RT   expression of green fluorescent protein fluorescence exclusively in
RT   germ cells: A useful model to monitor germ cells in a live
RT   vertebrate.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:2544-2549(2001).
DR   EMBL; AB063484; BAB61047.1; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; UNKNOWN_1.
KW   ATP-binding; Helicase.
SQ   SEQUENCE   617 AA;  66737 MW;  C1B8A3462760BA50 CRC64;

Alignment Scores:
Pred. No.:               9.27e-140     Length:        617
Score:                   1899.50       Matches:       386
Percent Similarity:      74.45%        Conservative:  86
Best Local Similarity:   60.88%        Mismatches:    129
Query Match:             48.04%        Indels:        33
DB:                      13            Gaps:          13

US-09-714-865-15 (1-2172) x Q90ZF6 (1-617)

Qy        304 GAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCA 363
Db          6 GluGluGluThrAlaProSerPheAlaProValSerSerThrAspAlaAla------Pro 23

Qy        364 ACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGAT---GGAAATAATTCAGAA 420
Db         24 GlnArgSer---------SerTrpAsnGlyGlyAlaArgAspSerGlyAsnAspGlyAsp 40

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC---CGAGGTTGCCGTGGAGGA 477
Db         41 SerTrpAsnArgSerAsnArgGlyArgGlyGlySerAlaGlyArgGlyGlyArgGlyGly 60
```

```
Qy        478 TTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGT 537
Db         61 ArgGlyArgGlyPheGlyArgSerAspGlnAsp--------------GluLeuAsnGly 75

Qy        538 GGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAA 597
Db         76 Gly---GlyGlyAspSerGluAsnGlyPheArgGlyArgGlyArgGlyGlyArgGlyGly 94

Qy        598 AGCAGAAGTGGCAGTGGAAGTGAACGA---------GGTGGTTACAAAGGTTTAAATGAA 648
Db         95 PheArgSerGly---GlyGlyGluArgGlyArgGlyGlyGlyTyrArgGlyArgAspGlu 113

Qy        649 GAAGTAATA---ACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
Db        114 AspValPheAlaAlaGlyAspGlyArgGly---------------AlaGluAsnSerAsp 128

Qy        706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAG 765
Db        129 AlaAlaAspProGluArgProLysValThrTyrIleProProSerLeuProGluAspGlu 148

Qy        766 GACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTT 825
Db        149 AspSerIlePheSerHisTyrLysMetGlyIleAsnPheAspLysTyrAspAspIleLeu 168

Qy        826 GTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTC 885
Db        169 ValAspValSerGlyThrAsnLeuProAlaAlaIleMetThrPheGluGluAlaLysLeu 188

Qy        886 TGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAA 945
Db        189 CysGluSerLeuGluAsnAsnIleSerArgSerGlyTyrValLysProThrProValGln 208

Qy        946 AAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGG 1005
Db        209 LysTyrGlyLeuProIleIleSerAlaGlyArgAspLeuMetAlaCysAlaGlnThrGly 228

Qy       1006 TCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATA 1065
Db        229 SerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnGlnLeuMetAlaAspGlyVal 248

Qy       1066 ACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGA 1125
Db        249 AlaAlaSerArgPheSerGluIleGlnGluProGluAlaValIleValAlaProThrArg 268

Qy       1126 GAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGA 1185
Db        269 GluLeuIleAsnGlnIleTyrGlnGluAlaArgLysPheSerPheGlyThrCysValArg 288

Qy       1186 GCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGC 1245
Db        289 ProValValValTyrGlyGlyValAsnThrGlyTyrGlnMetArgGluIleGluLysGly 308

Qy       1246 TGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATT 1305
Db        309 CysAsnValLeuCysGlyThrProGlyArgLeuLeuAspMetIleGlyArgGlyLysVal 328

Qy       1306 GGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGT 1365
Db        329 GlyLeuSerLysValArgHisLeuValLeuAspGluAlaAspArgMetLeuAspMetGly 348

Qy       1366 TTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGC 1425
Db        349 PheGluProAspMetArgArgLeuValGlySerProGlyMetProSerLysGluGluArg 368

Qy       1426 CAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTT 1485
Db        369 GlnThrLeuMetPheSerAlaThrPheProGluAspIleGlnArgLeuAlaAlaAspPhe 388

Qy       1486 TTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTT 1545
Db        389 LeuLysValAspTyrLeuPheValAlaValGlyValValGlyGlyAlaCysThrAspVal 408

Qy       1546 CAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTG 1605
Db        409 GluGlnThrPheLeuGlnValThrLysPheAsnLysArgGluGlnLeuLeuAspLeuLeu 428

Qy       1606 CGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTT 1665
Db        429 ArgThrIleGlySerGluArgThrMetValPheValGluThrLysArgGlnAlaAspPhe 448

Qy       1666 ACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAA 1725
Db        449 IleAlaAlaPheLeuCysGlnGluLysValProThrThrSerIleHisGlyAspArgGlu 468

Qy       1726 CAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTT 1785
Db        469 GlnArgGluArgGluLysAlaLeuAlaAspPheArgSerGlyLysCysProValLeuVal 488

Qy       1786 GCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTT 1845
Db        489 AlaThrSerValAlaSerArgGlyLeuAspIleProAspValGlnHisValValAsnPhe 508

Qy       1846 GATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGG 1905
Db        509 AspLeuProAsnThrIleAspAspTyrValHisArgIleGlyArgThrGlyArgCysGly 528

Qy       1906 AATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCT 1965
Db        529 AsnThrGlyArgAlaValSerPheTyrAspProAspValAspSerGlnLeuAlaArgSer 548

Qy       1966 CTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCC 2025
Db        549 LeuValGlyIleLeuAlaLysAlaGlnGlnGluValProSerTrpLeuGluGluSerAla 568

Qy       2026 TTTAGTACATACATTCCTGGCTTCAGTGGTAGTACA------------AGAGGAAACGTG 2073
Db        569 Phe---------------GlyAlaHisGlySerAlaAlaPheAsnProSerGlyArgThr 583

Qy       2074 TTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCT 2133
Db        584 PheAlaSerThrAspSerArgLysGlyGlySerPheGlnAspSerSerValLysThrGln 603

Qy       2134 TCACGAGCT---CCCAATCCAGTAGATGATGAGTCATGGGAT 2172
Db        604 ProAlaAlaProProAlaAlaAlaAspGluAspAspTrpGlu 617
```

```
RESULT 7
Q9PT10
ID   Q9PT10      PRELIMINARY;      PRT;   647 AA.
AC   Q9PT10;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Vasa.
GN   VAS.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=20160526; PubMed=10694742;
RA   Yoshizaki G., Sakatani S., Tominaga H., Takeuchi T.;
RT   "Cloning and characterization of a vasa-like gene in rainbow trout and
RT   its expression in the germ cell lineage.";
RL   Mol. Reprod. Dev. 55:364-371(2000).
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AB032566; BAA88059.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
```

```
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   647 AA;  68525 MW;  88F25AE7A68F51C2 CRC64;

Alignment Scores:
Pred. No.:              1.27e-138     Length:        647
Score:                  1885.00       Matches:       395
Percent Similarity:     68.32%        Conservative:  86
Best Local Similarity:  56.11%        Mismatches:    141
Query Match:            47.67%        Indels:        82
DB:                     13            Gaps:          15

US-09-714-865-15 (1-2172) x Q9PT10 (1-647)

Qy     79 TCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCAGAAATGGATGAT 138
Db     20 SerGlyGlyGlnSerSerSerPheGlyArgProThrAspLysValSerSerTrpAsnSer 39

Qy    139 GGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGA 198
Db     40 Gly-----------------------CysGlyGlyPheGlyGlyGlyArgGlyArgGly 51

Qy    199 AACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGA 258
Db     52 SerArg------------------------------------------GlyGlyGlyGly 57

Qy    259 GTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGC 318
Db     58 AspPheLysSerPheSerSerGlyAlaGlyGluAsnGlyAsn---GluAspLysAspSer 76

Qy    319 TCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGG 378
Db     77 Ser------Trp------------------------------------------------ 78

Qy    379 TTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGA 438
Db     79 ---------------------------------------AsnSerGlyGlyGlyPheArg 85

Qy    439 AGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAAT 498
Db     86 GlyArgGlyArgGlyGlyGlyArgGlySerArgGlyGly---------GlyIleArgAsn 102

Qy    499 AATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGA 558
Db    103 GlyAspAspGlyTyrAsp-----------------GlyGlyPheGlyGlySerGln--- 115

Qy    559 CCAGTATTAAGTGGCACAGGT------AATGGTGATACTTCTCAAAGCAGAAGTGGCAGT 612
Db    116 ---------GlyGlyArgGlyGlyArgGlyGlyPheArgSerGlyGlyGlyAspGlyGly 132

Qy    613 GGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACA---GGCTCTGGA 669
Db    133 GlyPheGlyGlyGlyGlyTyrArgGlyArgAspGluGluIlePheSerLysGlySerThr 152

Qy    670 AAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAA 729
Db    153 MetAspGly-----------------GlyGlyAspGlyGlyAsnProGlyProProLys 166

Qy    730 GTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAG 789
Db    167 ValThrTyrValProGluAlaLeuCysGluLysGluSerSerIlePheAlaHisTyrGlu 186

Qy    790 ACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCA 849
Db    187 SerGlyIleAsnPheAspLysTyrAspAspIleLeuValAspValSerGlySerAsnPro 206

Qy    850 CCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATT 909
Db    207 ProLysAlaIleMetGlyPheGluGluAlaAlaLeuCysGluSerLeuAsnArgAsnIle 226

Qy    910 GCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTT 969
Db    227 SerLysSerGlyTyrLysLysProThrProValGlnLysHisGlyIleProIleIleAla 246

Qy    970 GCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTC 1029
Db    247 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 266

Qy   1030 CTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTG 1089
Db    267 LeuProIleLeuGlnGlnLeuMetValAspGlyValAlaAlaSerGlnPheSerGluIle 286

Qy   1090 CAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149
Db    287 GlnGluProGluValIleIleValAlaProThrArgGluLeuIleAsnGlnIleTyrMet 306

Qy   1150 GAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACC 1209
Db    307 GluAlaArgLysPheAlaHisGlyThrCysValArgProValValValTyrGlyGlyIle 326

Qy   1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269
Db    327 SerThrGlyHisThrIleArgGluIleLeuLysGlyCysAsnValLeuCysAlaThrPro 346

Qy   1270 GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA 1329
Db    347 GlyArgLeuMetAspIleIleGlyArgGlyLysIleGlyLeuSerLysLeuArgTyrLeu 366

Qy   1330 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTA 1389
Db    367 ValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProAlaMetArgLysLeu 386

Qy   1390 ATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACT 1449
Db    387 ValGlySerProGlyIleProAlaLysGluAspArgGlnThrLeuMetPheSerAlaThr 406

Qy   1450 TTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTT 1509
Db    407 TyrProGluAspIleGlnLysLeuAlaAlaAspPheLeuLysLysAspTyrLeuPheLeu 426

Qy   1510 GCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGC 1569
Db    427 AlaValGlyValValGlyGlyAlaCysSerAspValGluGlnValValValGlnValThr 446

Qy   1570 CAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACT 1629
Db    447 LysPheSerLysArgAspGlnLeuLeuGluValLeuLysThrThrGlySerGluArgThr 466

Qy   1630 ATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAA 1689
Db    467 MetValPheValGluThrLysArgGlnAlaAspPheIleAlaThrPheLeuCysArgGlu 486

Qy   1690 AAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTT 1749
Db    487 LysValAsnThrThrSerIleHisGlyAspArgGluGlnArgGluArgGluGlnAlaLeu 506

Qy   1750 GGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGG 1809
Db    507 GlyAspPheArgSerGlyArgCysProValLeuValAlaThrSerValAlaAlaArgGly 526

Qy   1810 CTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAA 1869
Db    527 LeuAspIleLysAspValGlnHisIleValAsnPheAspLeuProAsnAsnIleAspGlu 546

Qy   1870 TATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTT 1929
Db    547 TyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAlaValCysPhe 566

Qy   1930 TTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCT 1989
Db    567 PheAspProGlyAlaAspGlyAsnLeuAlaArgSerLeuValLysValLeuSerGlyAla 586

Qy   1990 CAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTC 2049
```

```
Db        587 GlnGlnGluValProLysTrpLeuGluGluAlaAlaPheSerAlaHis------GlyThr 604

Qy       2050 AGTGGT---AGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGC 2106
              :::|||   :::       |||       |||||||||   |||:::||||||||   |||
Db        605 ThrGlyPheAsnProGlyGlyArgThrPheAlaSerThrAspSerArgLysGlyGlySer 624

Qy       2107 ACTTTGAACACAGCTGGGTTTTCT------TCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
                 |||         |||   |||       ::::::         :::      ||||||
Db        625 ---LeuGlnArgAspGlyAlaSerHisLeuAlaAlaAlaLeuSerSerGlyAlaAspAsp 643

Qy       2161 GAGTCATGGGAT 2172
              :::   |||:::
Db        644 AspGluTrpGlu 647

RESULT 8
Q8QHL6
ID   Q8QHL6      PRELIMINARY;      PRT;    621 AA.
AC   Q8QHL6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa short form.
GN   VAS-S.
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21664110; PubMed=11804791;
RA   Kobayashi T., Kajiura-Kobayashi H., Nagahama Y.;
RT   "Two isoforms of vasa homologs in a teleost fish: their differential
RT   expression during germ cell differentiation.";
RL   Mech. Dev. 111:167-171(2002).
DR   EMBL; AB051835; BAB56110.1; -.
SQ   SEQUENCE   621 AA;  67651 MW;  F0A828FDC5308FAA CRC64;

Alignment Scores:
Pred. No.:               1.09e-137     Length:         621
Score:                   1873.00       Matches:        382
Percent Similarity:      67.09%        Conservative:   89
Best Local Similarity:   54.42%        Mismatches:     123
Query Match:             47.37%        Indels:         108
DB:                      13            Gaps:           13

US-09-714-865-15 (1-2172) x Q8QHL6 (1-621)

Qy        109 ACTCCAGCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAA 168
              |||   |||   :::|||:::            |||   ::::::   |||
Db         14 ThrCysAlaLeuThrSerAspThrSerThrGlyGlyThrGlnGlyAsp------------ 29

Qy        169 AGTGGATTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGA 228
                                     |||   |||   |||   ||||||
Db         30 -----------------------PheTrpAsnSerAspHisGlyGlu------------ 37

Qy        229 GATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTT 288
                          ||||||   |||   |||:::      |||      ||||||
Db         38 ------------PheGlnGlyGlyArgGly---GlyArgGlyArgGlyArgGlyGlyPhe 52

Qy        289 TCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGAC 348
                 ||||||   |||   |||||||||
Db         53 LysAsnSerPhePheSerAspGlyAspGly------------------------------ 62

Qy        349 TGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGC---------TAT 399
                 ::::::|||                  ||||||:::||||||||
Db         63 ---GlnAsnAsnThrAlaGluArgSerSerPheSerArgArgGlyGlyArgAlaAsnGlu 81

Qy        400 CGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC 459
                 ||||||:::|||          |||      |||||||||      |||
Db         82 GluAspGlyAspAsnLysAsnValSerArgGlyArgArgGlyArgGlyGlyGlyGlu 101

Qy        460 CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAA 519
              :::|||   |||   |||||||||
Db        102 GlnGlyGlyArgSerGlyPheGly------------------------------------- 109

Qy        520 TGTATGCAGCGCACTGGTGGCCTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGT 579

Db        109 ------------------------------------------------------------ 109

Qy        580 AATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGT 639
                                                           |||||||||:::|||
Db        110 ---------------------------------------------GlyGlyTyrArgGly 114

Qy        640 TTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAA--------- 690
                 :::|||                           |||:::      |||:::
Db        115 LysAspGlu---------------------LysSerLeuLeuLysAlaArgGlyTyrPro 127

Qy        691 GCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCT 750
              ||||||      ::::::||||||         |||:::|||||||||:::|||||||||
Db        128 AlaGluAsnAlaAspAlaSerAspArg------ProArgValThrTyrValProProPro 145

Qy        751 CCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAA 810
              |||   |||||||||:::|||||||||:::||||||:::|||:::|||||||||||||||
Db        146 ProSerGluAspGluGluSerIlePheSerHisTyrGluSerGlyValAsnPheAspLys 165

Qy        811 TACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTT 870
              ||||||   |||||||||:::|||||||||      :::   ||||||||||||:::||||||
Db        166 TyrAspGluIleLeuValAspValSerGlyThrAsnProProProAlaIleMetThrPhe 185

Qy        871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
              :::||||||   ||||||::::::|||         |||::::::|||:::||||||   |||
Db        186 AspGluAlaAlaLeuCysGluSerLeuLysArgAsnValSerLysSerGlyTyrValLys 205

Qy        931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
                 |||||||||||||||:::   ||||||||||||   |||||||||||||||||||||
Db        206 ProThrProValGlnLysHisGlyIleProIleIleSerAlaGlyArgAspLeuMetAla 225

Qy        991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG 1050
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||   :::
Db        226 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnGlnLeu 245

Qy       1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATT 1110
              |||   ||||||         ||||||   |||   |||:::|||||||||:::   ||||||
Db        246 MetAlaAspGlyAlaAlaAlaSerCysPheSerGluMetGlnGluProAspAlaIleIle 265

Qy       1111 GTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTT 1170
              |||||||||||||||||||:::||||||||||||||||||||||||||||||::::::
Db        266 ValAlaProThrArgGluLeuIleAsnGlnIleTyrLeuGluAlaArgLysPheAlaTyr 285

Qy       1171 GGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGA 1230
              ||||||||||||||   ||||||:::|||||||||            ||||||   ||||||
Db        286 GlyThrCysValArgProValValValTyrGlyGlyValSerThrGlyHisGlnIleArg 305

Qy       1231 CAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATA 1290
                 :::::::::|||||||||:::||||||   ||||||||||||:::|||:::|||
Db        306 AspLeuLeuArgGlyCysAsnValLeuCysGlyThrProGlyArgLeuLeuAspMetIle 325

Qy       1291 GGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGC 1350
              |||:::   |||:::||||||   :::::::::||||||||||||||||||||||||||||
Db        326 GlyArgGlyLysValGlyLeuThrLysValArgTyrLeuValLeuAspGluAlaAspArg 345

Qy       1351 ATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCA 1410
              ||||||||||||||||||   |||:::|||::::::|||:::         ||||||||||||
Db        346 MetLeuAspMetGlyPheGluProAspMetArgArgLeuValGlySerProGlyMetPro 365

Qy       1411 TCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGG 1470
              |||||||||   ||||||||||||||||||||||||||||||||||||:::|||||||||
Db        366 SerLysGluAsnArgGlnThrLeuMetPheSerAlaThrPheProGluAspIleGlnArg 385

Qy       1471 TTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGA 1530
              |||||||||:::|||||||||::::::|||||||||:::|||||||||   |||||||||
Db        386 LeuAlaAlaAspPheLeuLysThrAspTyrLeuPheLeuAlaValGlyIleValGlyGly 405

Qy       1531 GCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAG 1590
```

```
                 |||||  ||||||:::||||||   :::||||||   :::|||:::|||||||||:::
Db         406 AlaCysSerAspValGluGlnThrPheValGlnValThrLysPheAlaLysArgGluGln 425

Qy        1591 CTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAG 1650
                 |||:::::::::|||:::::    |||   ||||||||||||||||||||||||||||
Db         426 LeuIleAspLeuLeuLysSerThrGlySerGluArgThrMetValPheValGluThrLys 445

Qy        1651 AAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATC 1710
               ::::::||||||||   ||||||  ||||||||||||||            ||||||||||||
Db         446 ArgGlnAlaAspPheIleAlaThrIleLeuCysGlnGluLysPheProThrThrSerIle 465

Qy        1711 CATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTCGCTTTGGAAAG 1770
               |||||||||||||||||     :::|||||||||||||||||||||||||||     ||||||
Db         466 HisGlyAspArgGluGlnTrpGlnArgGluGlnAlaLeuGlyAspPheArgSerGlyLys 485

Qy        1771 TGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAA 1830
               |||   ||||||||||||||||||||   |||||||||||||||||     :::||||||
Db         486 CysSerValLeuValAlaThrSerValGlyAlaArgGlyLeuAspIleProAspValGln 505

Qy        1831 CATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGT 1890
                  |||:::|||||||||||||||:::    |||||||||||||||||||||||||||||
Db         506 PheValValAsnPheAspLeuProAsnAsnIleAspGluTyrValHisArgIleGlyArg 525

Qy        1891 ACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAAC 1950
               ||||||||||||||||||||||||||||||:::||||||:::|||       :::|||:::
Db         526 ThrGlyArgCysGlyAsnThrGlyArgAlaValSerPheTyrAspProGlyAlaAspSer 545

Qy        1951 CATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGG 2010
                  ||||||:::   ||||||    :::|||:::   |||||||||:::||||||:::|||
Db         546 GluLeuAlaArgSerLeuValThrIleLeuSerLysAlaGlnGlnGluValProSerTrp 565

Qy        2011 TTGGAAGAAATTGCCTTTAGTACATACATTCCT---GGCTTCAGTGGTAGTACAAGAGGA 2067
               |||||||||  |||||||||    :::    |||   ||||||:::            :::
Db         566 LeuGluGluPheAlaPheSerGlyHisSerProSerGlyPheAsnProProArgLysHis 585

Qy        2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
                     |||||||||   |||    ||||||||                     ||||||
Db         586 ------PheAlaSerThrAspLeuArgLysGlyProGln-----------GlyGlyPhe 599

Qy        2128 ---------------------TCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCA 2166
                                    |||       |||          ||||||||
Db         600 PheGlnAspGlySerValThrSerGlnLeuAlaAlaGlnThrAlaAlaAspAspGluGlu 619

Qy        2167 TGGGAT 2172
               |||:::
Db         620 TrpGlu 621
```

RESULT 9
Q9DEG3

```
ID   Q9DEG3      PRELIMINARY;      PRT;   645 AA.
AC   Q9DEG3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa.
GN   VAS.
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20544961; PubMed=11091081;
RA   Kobayashi T., Kajiura-Kobayashi H., Nagahama Y.;
RT   "Differential expression of vasa homologue gene in the germ cells
RT   during oogenesis and spermatogenesis in a teleost fish, tilapia,
RT   Oreochromis niloticus.";
RL   Mech. Dev. 99:139-142(2000).
DR   EMBL; AB032467; BAB19807.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR000362; Fumarate_lyase.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000685; RuBisCO_large.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS00163; FUMARATE_LYASES; UNKNOWN_1.
DR   PROSITE; PS00157; RUBISCO_LARGE; UNKNOWN_1.
KW   ATP-binding; Helicase.
SQ   SEQUENCE   645 AA;  70072 MW;  4C945536A9BF21DA CRC64;

Alignment Scores:
Pred. No.:              4.63e-137     Length:         645
Score:                  1865.00       Matches:        389
Percent Similarity:     68.80%        Conservative:   94
Best Local Similarity:  55.41%        Mismatches:     135
Query Match:            47.17%        Indels:         84
DB:                     13            Gaps:           18

US-09-714-865-15 (1-2172) x Q9DEG3 (1-645)

Qy         109 ACTCCAGCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAA 168
                |||   |||   :::|||:::                |||   ::::::   |||
Db          14 ThrCysAlaLeuThrSerAspThrSerThrGlyGlyThrGlnGlyAsp------------ 29

Qy         169 AGTGGATTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGA 228
                                       |||   |||   |||   |||||
Db          30 ------------------------PheTrpAsnSerAspHisGlyGlu------------ 37

Qy         229 GATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTT 288
                           ||||||   |||   |||:::         |||      ||||||
Db          38 ------------PheGlnGlyGlyArgGly---GlyArgGlyArgGlyArgGlyGlyPhe 52

Qy         289 TCAAACAGCAGGTTT---------GAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCT 339
                 ||||||   |||         |||||||||   :::   |||
Db          53 LysAsnSerPhePheSerGlyAlaGluAspGlyAlaAsnAspGly--------------- 67

Qy         340 AGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTAT 399
                     |||   ::::::|||               ||||||:::|||||||||
Db          68 AspGlyAspGlyGlnAsnAsnThrAlaGluArgSerSerPheSerArgArgGlyGly--- 86

Qy         400 CGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC 459
                                                          |||||||||
Db          87 ------------------------------------------ArgGlyArgGlyLys--- 91

Qy         460 CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAA 519
                              |||||||||   :::   ::::::|||   |||   :::|||
Db          92 ---------------GlyPheGly---ArgThrAspHisSerAspPheAspAlaAsnGlu 105

Qy         520 TGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGT 579
                                 :::   ||||||             |||   |||
Db         106 GluAspGlyAspAsnLysAsnValSerArgGlyArgArg---------ArgGlyArgGly 122

Qy         580 AATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGT 639
                  |||:::         |||||||||   |||            |||||||||:::|||
Db         123 GlyGlyGluGlnGly---GlyArgSerGlyPheGly---------GlyGlyTyrArgGly 138

Qy         640 TTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAA--------- 690
                  :::|||                     |||:::      |||:::
Db         139 LysAspGlu--------------------LysSerLeuLeuLysAlaArgGlyTyrPro 151

Qy         691 GCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCT 750
                ||||||      ::::::||||||         |||:::|||||||||:::|||||||||
Db         152 AlaGluAsnAlaAspAlaSerAspArg------ProArgValThrTyrValProProPro 169

Qy         751 CCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAA 810
                |||   |||||||||:::|||||||||:::||||||::::::|||:::||||||||||||
Db         170 ProSerGluAspGluGluSerIlePheSerHisTyrGluSerGlyValAsnPheAspLys 189

Qy         811 TACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTT 870
```

```
                  ||||||  |||||||||:::||||||||   :::   |||||||||||:::||||||
Db        190 TyrAspGluIleLeuValAspValSerGlyThrAsnProProProAlaIleMetThrPhe 209

Qy        871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
              :::||||||  ||||||::::::|||       |||::::::|||:::||||||   |||
Db        210 AspGluAlaAlaLeuCysGluSerLeuLysArgAsnValSerLysSerGlyTyrValLys 229

Qy        931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
                 |||||||||||||:::    ||||||||||||  |||||||||||||||||||||
Db        230 ProThrProValGlnLysHisGlyIleProIleIleSerAlaGlyArgAspLeuMetAla 249

Qy        991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||||     :::
Db        250 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnGlnLeu 269

Qy       1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATT 1110
              |||   ||||||       ||||||   |||   |||:::|||||||||:::   ||||||
Db        270 MetAlaAspGlyAlaAlaAlaSerCysPheSerGluMetGlnGluProAspAlaIleIle 289

Qy       1111 GTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTT 1170
              ||||||||||||||||||||:::||||||||||||||||||||||||||||||::::::
Db        290 ValAlaProThrArgGluLeuIleAsnGlnIleTyrLeuGluAlaArgLysPheAlaTyr 309

Qy       1171 GGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGA 1230
              ||||||||||||||||   ||||||:::|||||||||        ||||||   ||||||
Db        310 GlyThrCysValArgProValValValTyrGlyGlyValSerThrGlyHisGlnIleArg 329

Qy       1231 CAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATA 1290
                 :::::::::||||||||:::||||||   |||||||||||||||:::|||:::|||
Db        330 AspLeuLeuArgGlyCysAsnValLeuCysGlyThrProGlyArgLeuLeuAspMetIle 349

Qy       1291 GGCAAAGAAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGC 1350
              |||:::    |||:::||||||   :::::::::||||||||||||||||||||||||||
Db        350 GlyArgGlyLysValGlyLeuThrLysValArgTyrLeuValLeuAspGluAlaAspArg 369

Qy       1351 ATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCA 1410
              ||||||||||||||||||   |||:::|||::::::|||:::       |||||||||||
Db        370 MetLeuAspMetGlyPheGluProAspMetArgArgLeuValGlySerProGlyMetPro 389

Qy       1411 TCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGG 1470
              |||||||||   ||||||||||||||||||||||||||||||||||||:::|||||||||
Db        390 SerLysGluAsnArgGlnThrLeuMetPheSerAlaThrPheProGluAspIleGlnArg 409

Qy       1471 TTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGA 1530
              |||||||||:::||||||||||||||:::::::||||||||||:::|||||||||   |||||||||
Db        410 LeuAlaAlaAspPheLeuLysThrAspTyrLeuPheLeuAlaValGlyIleValGlyGly 429

Qy       1531 GCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAG 1590
              ||||||   ||||||:::||||||   :::||||||   :::|||:::|||||||||:::
Db        430 AlaCysSerAspValGluGlnThrPheValGlnValThrLysPheAlaLysArgGluGln 449

Qy       1591 CTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAG 1650
              |||:::::::::|||::::::   |||   ||||||||||||||||||||||||||||||
Db        450 LeuIleAspLeuLeuLysSerThrGlySerGluArgThrMetValPheValGluThrLys 469

Qy       1651 AAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATC 1710
              ::::::|||||||||   ||||||   |||||||||||||||      ||||||||||||
Db        470 ArgGlnAlaAspPheIleAlaThrIleLeuCysGlnGluLysPheProThrThrSerIle 489

Qy       1711 CATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAG 1770
              ||||||||||||||||||      :::|||||||||||||||||||||||||||   ||||||
Db        490 HisGlyAspArgGluGlnTrpGlnArgGluGlnAlaLeuGlyAspPheArgSerGlyLys 509

Qy       1771 TGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAA 1830
              |||   ||||||||||||||||||||   ||||||||||||||||||   :::||||||
Db        510 CysSerValLeuValAlaThrSerValGlyAlaArgGlyLeuAspIleProAspValGln 529

Qy       1831 CATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGT 1890
                 |||:::|||||||||||||||:::   ||||||||||||||||||||||||||||||
Db        530 PheValValAsnPheAspLeuProAsnAsnIleAspGluTyrValHisArgIleGlyArg 549

Qy       1891 ACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAAC 1950
              |||||||||||||||||||||||||||||:::||||||:::|||      :::|||:::

Db        550 ThrGlyArgCysGlyAsnThrGlyArgAlaValSerPheTyrAspProGlyAlaAspSer 569

Qy       1951 CATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGG 2010
                 ||||||:::   ||||||   :::|||:::   |||||||||:::||||||:::|||
Db        570 GluLeuAlaArgSerLeuValThrIleLeuSerLysAlaGlnGlnGluValProSerTrp 589

Qy       2011 TTGGAAGAAATTGCCTTTAGTACATACATTCCT---GGCTTCAGTGGTAGTACAAGAGGA 2067
                 |||||||||   |||||||||   :::   |||   ||||||:::         :::
Db        590 LeuGluGluPheAlaPheSerGlyHisSerProSerGlyPheAsnProProArgLysHis 609

Qy       2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
                          |||||||||   |||   |||||||||                 ||||||
Db        610 ------PheAlaSerThrAspLeuArgLysGlyProGln------------GlyGlyPhe 623

Qy       2128 --------------------TCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCA 2166
                                  |||         |||         |||||||||
Db        624 PheGlnAspGlySerValThrSerGlnLeuAlaAlaGlnThrAlaAlaAspAspGluGlu 643

Qy       2167 TGGGAT 2172
              |||:::
Db        644 TrpGlu 645
```

RESULT 10

```
Q9DGR8
ID   Q9DGR8      PRELIMINARY;      PRT;   662 AA.
AC   Q9DGR8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cvh.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20283536; PubMed=10821771;
RA   Tsunekawa N., Naito M., Sakai Y., Nishida T., Noce T.;
RT   "Isolation of chicken vasa homolog gene and tracing the origin of
RT   primordial germ cells.";
RL   Development 127:2741-2750(2000).
DR   EMBL; AB004836; BAB12337.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000685; RuBisCO_large.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS00157; RUBISCO_LARGE; UNKNOWN_1.
KW   ATP-binding; Helicase.
SQ   SEQUENCE   662 AA;  73328 MW;  BC244540133877FB CRC64;

Alignment Scores:
Pred. No.:               3.93e-129      Length:        662
Score:                   1763.50        Matches:       377
Percent Similarity:      64.19%         Conservative:  80
Best Local Similarity:   52.95%         Mismatches:    164
Query Match:             44.60%         Indels:        91
DB:                      13             Gaps:          11

US-09-714-865-15 (1-2172) x Q9DGR8 (1-662)

Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              :::|||||||||:::   |||:::                                |||
Db          2 GluGluAspTrpAspThrGluLeu--------------------------------Glu 10
```

```
Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
              ::::::      :::        |||   :::      :::        ::::::|||
Db         11 GlnGluAlaAlaAlaAlaSerGlnGlyArgSerGluGluGlnAlaTrpMetAlaAsnSer 30

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
                 :::  ||||||   |||
Db         31 GlyArgProAsnSerProSerLeuArg--------------------------------- 39

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
                    |||   :::|||                          :::|||   :::
Db         40 ------PheSerSerArgPro--------------------------SerSerProLeu 48

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
                 ||||||            :::         |||||||||   ::::::
Db         49 SerGlyPheProGlyArgProAsnSerProPhePheGlyPheSerGlnAsnLysGlySer 68

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
                 |||   :::   |||
Db         69 LeuGlyAlaAsnGluGlyLeu--------------------------------------- 75

Qy        367 CGGAACAGAGGGTTTTCCAAGAGA---------GGCGGCTATCGAGATGGAAATAATTCA 417
                 ||||||         :::         |||||||||                |||
Db         76 ---AsnArgSerLeuProValGlnHisAspIleGlyGlyTyrSerGlySerArgGluSer 94

Qy        418 ---------------GAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGA 462
                                      |||      |||   |||||||||
Db         95 ValValArgGlnAsnArgGluAspGlnProValThrArgPheGlyArgGlyArgSerSer 114

Qy        463 GGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGT 522
                 |||   |||                  |||   |||:::         |||:::
Db        115 GlySerArgAspPheGlnGluArgAsnSerAlaAsnAspProGlyMetGlnAspGlnGly 134

Qy        523 ATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAAT 582
                 :::|||      |||:::||||||      :::         |||         |||
Db        135 PheArgArgValProGlyIlePheGlyGlnSerLysCysPheAsnSerGluGluArgAsn 154

Qy        583 GGTGATACTTCTCAAAGC------------------AGAAGTGGCAGTGGAAGTGAACGA 624
                          |||          |||      |||
Db        155 SerProLeuArgGlySerProPheAlaProGlyGlyArgGlyAlaValGlyGlyProAla 174

Qy        625 GGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAG 684
              |||      ||||||   :::||||||:::   :::|||
Db        175 GlyValLeuLysGlyArgSerGluGluIleAspSerGly--------------------- 187

Qy        685 TCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCC 744
                                                 :::||||||||||||||||||:::|||
Db        188 ---------------------------------ArgGlyProLysValThrTyrValPro 196

Qy        745 CCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTC 804
              |||||||||||||||||||||   ||||||||||||   ||||||:::||||||||||||
Db        197 ProProProProGluAspGluGlnSerIlePheAlaCysTyrGlnSerGlyIleAsnPhe 216

Qy        805 GACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTG 864
              ||||||||||||         ||||||:::||||||   |||   |||      :::|||
Db        217 AspLysTyrAspGluCysAlaValGluMetSerGlyLeuAspProProAlaProLeuLeu 236

Qy        865 ACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTAT 924
                 |||||||||||||||         |||||||||      ||||||:::|||   ||||||
Db        237 AlaPheGluGluAlaAsnPheAlaGlnThrLeuArgLysAsnIleSerLysThrGlyTyr 256

Qy        925 ACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTG 984
              :::||||||||||||||||||||||:::|||||||||:::|||   |||||||||||||||
Db        257 SerLysLeuThrProValGlnLysHisSerIleProValIleGlnAlaGlyArgAspLeu 276

Qy        985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCT 1044
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        277 MetSerCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleValAsp 296

Qy       1045 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
                 ||||||   ||||||:::|||||||||      |||   :::   |||:::|||:::|||
Db        297 ArgMetMetLysAspGlyValThrAlaSer---PheProLysGlnGlnAspProGlnCys 315

Qy       1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164
              ||||||||||||||||||||||||||:::|||||||||:::||||||||||||||||||
Db        316 IleIleValAlaProThrArgGluLeuIleAsnGlnIlePheLeuGluAlaArgLysPhe 335

Qy       1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA 1224
                 :::|||||||||:::|||   |||||||||||||||||||||||   |||||||||
Db        336 ValTyrGlyThrCysIleArgProValValIleTyrGlyGlyThrGlnThrGlyHisSer 355

Qy       1225 ATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGAT 1284
              ||||||||||||:::||||||||||||||||||||||||||||||||||||:::|||
Db        356 IleArgGlnIleMetGlnGlyCysAsnIleLeuCysAlaThrProGlyArgLeuLeuAsp 375

Qy       1285 ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCT 1344
              ||||||   |||   ||||||   |||      ::::::||||||||||||||||||||||
Db        376 IleIleGluLysGlyLysIleSerLeuValGluValLysTyrLeuValLeuAspGluAla 395

Qy       1345 GATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
              ||||||||||||||||||||||||||      :::||||||||||||||||||   |||
Db        396 AspArgMetLeuAspMetGlyPheGlyLeuAspMetLysLysLeuIleSerTyrProGlu 415

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
              ||||||||||||::::::||||||||||||||||||||||||||||||||||||||||:::
Db        416 MetProSerLysAspArgArgGlnThrLeuMetPheSerAlaThrPheProGluGluVal 435

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
              ||||||||||||   ||||||||||||::::::|||:::|||:::      :::|||
Db        436 GlnArgLeuAlaGlyGluPheLeuLysThrAspTyrIlePheLeuValIleGlyAsnThr 455

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
                 |||||||||   ||||||||||||      :::|||||||||   :::   |||||||||
Db        456 CysGlyAlaCysSerAspValGlnGlnAsnIleLeuGlnValProArgLeuSerLysArg 475

Qy       1585 GAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAA 1644
              :::||||||:::|||||||||::::::   |||   ||||||||||||||||||||:::
Db        476 AspLysLeuIleGluIleLeuGlnSerThrGlyGlyGluArgThrMetValPheValAsp 495

Qy       1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACA 1704
              ||||||||||||||||||:::   |||   |||||||||||||||      :::   :::|||
Db        496 ThrLysLysLysAlaAspTyrLeuAlaAlaPheLeuCysGlnGluAsnLeuProSerThr 515

Qy       1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
              ||||||||||||||||||||||||||||||||||||||||      ||||||||||
Db        516 SerIleHisGlyAspArgGluGlnArgGluArgGluIleAlaLeuArgAspPheArgSer 535

Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824
              |||||||||   :::||||||||||||||||||||||:::||||||||||||||||||
Db        536 GlyLysCysGlnIleLeuValAlaThrSerValAlaSerArgGlyLeuAspIleGluAsn 555

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884
              |||||||||||||||||||||||||||:::||||||::::::|||||||||||||||
Db        556 ValGlnHisValIleAsnPheAspLeuProAsnThrIleGluAspTyrValHisArgIle 575

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCG 1944
              ||||||||||||||||||||||||||||:::|||:::||||||||||||      :::|||
Db        576 GlyArgThrGlyArgCysGlyAsnThrGlyLysAlaValSerPhePheAspAspGlnSer 595

Qy       1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
              |||   ||||||   |||   |||:::|||||||||:::         ||||||
Db        596 AspGlyHisLeuValGlnSerLeuLeuLysValLeuSerArgThrGlnGln--------- 612

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGA 2064
                          |||         |||                        |||
Db        613 ------------GluPheGlnPheGly------------GlyArgMetAlaValGlnArg 624

Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGC 2100
                 |||:::   ||||||         ||||||
Db        625 ThrAsnIleValAlaSerThrTrpCysProLysGly 636
```

RESULT 11
Q9GNP1
ID Q9GNP1 PRELIMINARY; PRT; 770 AA.
AC Q9GNP1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

```
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa homolog.
GN   CSDEAD1B(CSVHB).
OS   Ciona savignyi.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=51511;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=20130953; PubMed=10664149;
RA   Fujimura M., Takamura K.;
RT   "Characterization of an ascidian DEAD-box gene, Ci-DEAD1: specific
RT   expression in the germ cells and its mRNA localization in the
RT   posterior-most blastomeres in early embryos.";
RL   Dev. Genes Evol. 210:64-72(2000).
DR   EMBL; AB047803; BAB12217.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF00098; zf-CCHC; 6.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 6.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase.
SQ   SEQUENCE   770 AA;  82032 MW;  5C6D2A2D8C9CDD58 CRC64;

Alignment Scores:
Pred. No.:               5.25e-114     Length:         770
Score:                   1570.00       Matches:        364
Percent Similarity:      57.93%        Conservative:   100
Best Local Similarity:   45.44%        Mismatches:     229
Query Match:             39.71%        Indels:         108
DB:                      5             Gaps:           21

US-09-714-865-15 (1-2172) x Q9GNP1 (1-770)

Qy        1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              |||   |||:::||||||:::|||        :::      :::|||:::
Db        1 MetSerAspAspAspTrpAspAlaProSerSerGlyProValSerThr------------ 16

Qy       61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
                        :::   |||     |||   :::|||        |||     :::
Db       17 ------------GlyPheGlyGly------GlySerSerPheGlyGluThrIleSerLys 30

Qy      121 TCATCAGAAATGGATGATGGA------CCTTCTCGAAGAGATCATTTCATGAAAAGTGGA 174
                                 |||               |||   |||:::         :::
Db       31 GlyGlyGlyArgGlyArgGlyLysPheAlaAspArgGlyAspTyrAsnGlnProAsnAsp 50

Qy      175 TTTGCCTCTGGGCGGAATTTTGGAAACAGAGAT--------------------------- 207
              :::       |||      ||||||     |||
Db       51 TyrGlyPheGlySerGlyPheGlyLysSerAspAspGlyGlyPheGlySerLysProAsn 70

Qy      208 ------GCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGA--- 258
                  |||:::   |||      |||:::|||          ||||||||||||
Db       71 SerGlyPheGlyLysSerAsnPheAspAspAspThrGlyPheGlyGlyGlyPheGlySer 90

Qy      259 ---GTTGGAAAGAGTTTTGGAAAC---AGAGGTTTTTCAAACAGCAGG------TTTGAA 306
                   |||      ||||||:::      ||||||   |||   |||:::      |||:::
Db       91 SerSerGlyGlyGlyPheGlyAspThrArgGlySerSerArgSerLysGlyCysPheLys 110

Qy      307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
                |||:::                                   |||   :::|||
Db      111 CysGlyGluGluGly------------HisMetSerArgGluCysProGlnGlyGlyGly 126

Qy      367 CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCA 426
                :::||||||         :::|||   ::::::   |||               |||
Db      127 GlySerArgGly---------LysGlyCysPheLysCysGlyGluGluGlyHisMetSer 143

Qy      427 GGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTA 486
                :::||||||   |||               |||   ||||||   |||
Db      144 ArgGluCysProLysGlyGlyGlyGlyGlyGlyGlyGlyGlyArgGlyCysPheLysCys 163

Qy      487 GGAAGT------------------------CCAAATAATGACTTAGACCCAGACGAATGT 522
              |||                                 |||                   |||
Db      164 GlyGluGluGlyHisMetSerArgGluCysProLysGlyGlyAspSerGlyPheGluGly 183

Qy      523 ATGCAGCGCACTGGTGGCCTTTTT------------------GGTTCTAGAAGACCAGTA 564
                    |||:::   |||   |||                               |||
Db      184 ArgSerArgSerLysGlyCysPheLysCysGlyGluGluGlyHisMetSerArgGluCys 203

Qy      565 TTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGA 624
                 |||   |||   |||                           |||||||||      :::
Db      204 ProGlnGlyGlyGlyGlyGly------------------ArgGlySerGlyCysPheLys 217

Qy      625 GGTGGTTACAAAGGT---TTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGG 681
                 |||   :::|||   ::::::   |||         |||   |||
Db      218 CysGlyGluGluGlyHisMetSerArgGluCysProGlnGlyGlyGlyGlyGlyArgGly 237

Qy      682 -----------------------------------------------------AAGTCA 687
                                                                      :::
Db      238 SerGlyCysPheLysCysGlyGluGluGlyHisMetSerArgGluCysProArgAsnThr 257

Qy      688 GAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCT 747
                 ||||||||||||   ||||||         |||   :::   ||||||||||||
Db      258 SerGlyGluGlyGlyGluLysSerAspArg------ProProIle---TyrIleProPro 274

Qy      748 CCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGAC 807
              ||||||||||||||||||         :::||||||      |||   ||||||||||||
Db      275 ProProProGluAspGluValGluMetPheAlaSerMetGlnArgGlyIleAsnPheGly 294

Qy      808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
              |||||||||   |||   |||||||||||||||   :::||||||   :::|||   |||
Db      295 LysTyrAspAlaIleProValGluValSerGlyValAsnAlaProLysSerIleProThr 314

Qy      868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
              ||||||   |||   |||   :::|||:::      |||:::   :::|||   |||
Db      315 PheGluValAlaGlyLeuProGluThrValLeuAlaAsnValLysArgAlaAsnTyrGlu 334

Qy      928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
              :::   |||||||||||||||||||||||||||||||   |||   ||||||||||||
Db      335 ArgProThrProValGlnLysTyrSerIleProIleIleAsnAlaAspArgAspLeuMet 354

Qy      988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              |||||||||||||||||||||||||||||||||||||||||||||:::|||
Db      355 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProValLeuThrLys 374

Qy     1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
              ::::::   :::|||:::   :::|||:::|||   |||   |||   |||   |||
Db      375 LeuIleThrAsnGlyLeuGlnSerSerGlnPheSerGluLysGlnThrProArgAlaIle 394

Qy     1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
              :::|||   |||||||||||||||:::   ||||||:::||||||||||||||||||
Db      395 ValValGlyProThrArgGluLeuIleTyrGlnIlePheLeuGluAlaArgLysPheSer 414

Qy     1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
                 ||||||   ||||||   ||||||   ||||||||||||   :::   |||   |||
Db      415 ArgGlyThrValValArgProValValAlaTyrGlyGlyThrSerMetAsnHisGlnIle 434

Qy     1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
              |||   :::   :::||||||:::||||||   |||||||||||||||||||||
Db      435 ArgAspLeuGlnArgGlyCysHisIleLeuIleAlaThrProGlyArgLeuMetAspPhe 454

Qy     1288 ATAGGCAAAGAAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
              |||   :::      :::||||||      :::::::::::::::||||||||||||||
Db      455 IleAsnArgGlyLeuValGlyLeuAspHisValGluPheValIleLeuAspGluAlaAsp 474

Qy     1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
              ||||||||||||||||||         |||::::::||||||   |||   |||||||||
Db      475 ArgMetLeuAspMetGlyPheGluThrGluIleArgLysLeuAlaSerSerProGlyMet 494
```

```
Qy       1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
              |||||||||        |||   |||||||||||||||||||||||||:::|||||||||
Db        495 ProSerLysSerAspArgHisThrLeuMetPheSerAlaThrPheProAspGluIleGln 514

Qy       1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGT 1527
              |||||||||   :::||||||:::   ::::::||||||:::   ||||||:::||||||
Db        515 ArgLeuAlaHisAspPheLeuArgGluAspPheLeuPheLeuThrValGlyArgValGly 534

Qy       1528 GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAA 1587
              |||||||||   ||||||   |||:::::::::||||||   |||      ||||||
Db        535 GlyAlaCysThrAspValThrGlnSerIleIleGlnValAspGlnAspAspLysArgAla 554

Qy       1588 AAGCTCGTTGAAATTCTGCGAAACATAGGGGAT------GAAAGAACTATGGTCTTTGTT 1641
              ||||||:::|||::::::   ::::::   :::         ||||||:::|||||||||
Db        555 LysLeuLeuGluLeuIleSerAspValAlaGluThrArgSerArgThrLeuValPheVal 574

Qy       1642 GAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACT 1701
              |||||||||:::   |||||||||   |||      |||   ||||||         |||
Db        575 GluThrLysArgGlyAlaAspPheLeuAlaCysMetLeuSerGlnGluGlyCysProThr 594

Qy       1702 ACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGC 1761
              |||||||||||||||||||||      |||:::||||||||||||||||||   ||||||:::
Db        595 ThrSerIleHisGlyAspArgLeuGlnGlnGluArgGluGlnAlaLeuArgAspPheLys 614

Qy       1762 TTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAA 1821
                          ||||||:::|||:::||||||||||||||||||||||||||||||||
Db        615 SerAlaValCysProIleLeuIleAlaThrSerValAlaAlaArgGlyLeuAspIlePro 634

Qy       1822 AATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGA 1881
                 |||:::||||||||||||:::|||:::|||      ||||||||||||||||||
Db        635 LysValGluHisValIleAsnTyrAspMetProLysGluIleAspGluTyrValHisArg 654

Qy       1882 ATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAA 1941
              |||||||||||||||||||||||||||   |||||||||   :::|||:::|||
Db        655 IleGlyArgThrGlyArgCysGlyAsnLeuGlyArgAlaThrThrPheTyrAspAsnAsn 674

Qy       1942 TCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTT 2001
                 |||         ||||||:::   |||||||||:::|||::::::|||||||||:::|||
Db        675 LysAspGlyGluLeuAlaArgSerLeuValLysIleLeuSerGluAlaGlnGlnGluVal 694

Qy       2002 CCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACA 2061
              |||   ||||||||||||   |||   |||                  ||||||:::
Db        695 ProGlyTrpLeuGluGluCysAlaGluSerAlaVal---------------GlySerSer 709

Qy       2062 AGAGGAAACGTG------TTTGCATCAGTTGATACCAGAAAG------------------ 2097
                 |||          |||      |||   ||||||
Db        710 PheGlyLysGluGlyGlyPheGlyGlyArgAspPheArgLysArgGlyGlyArgAsnAsp 729

Qy       2098 ------------------------GGCAAGAGCACTTTGAACACAGCT------------ 2121
                                      |||      :::   :::||||||
Db        730 ThrGlyPheGlyGlySerLysGlyGlyPheGlySerArgSerThrAlaAspCysAspTyr 749

Qy       2122 ------------GGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGG 2169
                          ||||||   ::::::   |||   :::      ||||||:::||||||
Db        750 AsnAspGlyGlyGlyPheGlyAsnThrAlaAlaValSerLysAspAspAspAspSerTrp 769

Qy       2170 GAT 2172
              |||
Db        770 Asp 770
```

RESULT 12
Q9GNP2

```
ID   Q9GNP2      PRELIMINARY;      PRT;    688 AA.
AC   Q9GNP2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vasa homolog.
GN   CSDEAD1A(CSVHA).
OS   Ciona savignyi.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=51511;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=20130953; PubMed=10664149;
RA   Fujimura M., Takamura K.;
RT   "Characterization of an ascidian DEAD-box gene, Ci-DEAD1: specific
RT   expression in the germ cells and its mRNA localization in the
RT   posterior-most blastomeres in early embryos.";
RL   Dev. Genes Evol. 210:64-72(2000).
DR   EMBL; AB047802; BAB12216.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF00098; zf-CCHC; 3.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 3.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase.
SQ   SEQUENCE   688 AA;  73744 MW;  7EE70CFE04A681E9 CRC64;
```

```
Alignment Scores:
Pred. No.:               1.79e-113      Length:          688
Score:                   1563.00        Matches:         358
Percent Similarity:      58.30%         Conservative:    88
Best Local Similarity:   46.80%         Mismatches:      201
Query Match:             39.53%         Indels:          118
DB:                      5              Gaps:            19
```

US-09-714-865-15 (1-2172) x Q9GNP2 (1-688)

```
Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATA 60
              |||   |||:::||||||:::|||
Db          1 MetSerAspAspAspTrpAspAla------------------------------------ 8

Qy         61 TTTGAGAAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCA 120
                                                                |||:::|||
Db          9 ---------------------------------------------------ProSerSer 11

Qy        121 TCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCC 180
                                ||||||                  :::   :::||||||
Db         12 ------------------GlyPro------------------ValSerThrGlyPheGly 19

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
                 |||   :::||||||                              :::   |||
Db         20 GlyGlySerSerPheGly------------------------------GluThrIleSer 29

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
                    ||||||   |||   ||||||   |||   :::||||||      |||
Db         30 LysGlyGlyGlyArgGlyArgGlyLys---PheAlaAspArgGlyAspTyrAsnGlnPro 48

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
                 :::   |||      |||||||||   :::
Db         49 AsnAspTyrGlyPheGlySerGlyPheGlyLysSerAspAspGlyGlyPheGlySerLys 68

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
              |||      |||   ||||||   |||                  |||
Db         69 Pro------AsnSerGlyPheGlyLysSerAsnPheAspAspAspThrGlyPheGlyGly 86

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGT---------TGCCGT 471
                 |||               ||||||   |||   ||||||               :::
Db         87 GlyPheGlySerSerSerGlyGlyGlyPheGlyAspThrArgGlySerSerArgSerLys 106

Qy        472 GGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGC 531
              |||   |||      |||         :::               ||||||   |||
Db        107 GlyCysPheLysCysGlyGluGluGlyHis------MetSerArgGluCysProGlnGly 124

Qy        532 ACTGGTGGCCTTTTTGGTTCT-------------------------------AGAAGA 558
                 ||||||      ||||||                                |||
```

```
Db          125 GlyGlyGlyGlyArgGlySerGlyCysPheLysCysGlyGluGluGlyHisMetSerArg 144

Qy          559 CCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT 618
Db          145 GluCysProGlnGlyGlyGlyGlyGly-----------------ArgGlySerGlyCys 158

Qy          619 GAACGAGGTGGTTACAAAGGT---TTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAAT 675
Db          159 PheLysCysGlyGluGluGlyHisMetSerArgGluCys-------------------- 171

Qy          676 TCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACC 735
Db          172 ProArgAsnThrSerGlyGluGlyGlyGluLysSerAspArg------ProProIle--- 188

Qy          736 TACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGC 795
Db          189 TyrIleProProProProProGluAspGluValGluMetPheAlaSerMetGlnArgGly 208

Qy          796 ATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCA 855
Db          209 IleAsnPheGlyLysTyrAspAlaIleProValGluValSerGlyValAsnAlaProLys 228

Qy          856 GCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAA 915
Db          229 SerIleProThrPheGluValAlaGlyLeuProGluThrValLeuAlaAsnValLysArg 248

Qy          916 GCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGA 975
Db          249 AlaAsnTyrGluArgProThrProValGlnLysTyrSerIleProIleIleAsnAlaAsp 268

Qy          976 CGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCA 1035
Db          269 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuPro 288

Qy         1036 ATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA 1095
Db          289 ValLeuThrLysLeuIleThrAsnGlyLeuGlnSerSerGlnPheSerGluLysGlnThr 308

Qy         1096 CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCC 1155
Db          309 ProArgAlaIleValValGlyProThrArgGluLeuIleTyrGlnIlePheLeuGluAla 328

Qy         1156 AGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTG 1215
Db          329 ArgLysPheSerArgGlyThrValValArgProValValAlaTyrGlyGlyThrSerMet 348

Qy         1216 GGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGA 1275
Db          349 AsnHisGlnIleArgAspLeuGlnArgGlyCysHisIleLeuIleAlaThrProGlyArg 368

Qy         1276 CTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTG 1335
Db          369 LeuMetAspPheIleAsnArgGlyLeuValGlyLeuAspHisValGluPheValIleLeu 388

Qy         1336 GATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT 1395
Db          389 AspGluAlaAspArgMetLeuAspMetGlyPheGluThrGluIleArgLysLeuAlaSer 408

Qy         1396 TGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCA 1455
Db          409 SerProGlyMetProSerLysSerAspArgHisThrLeuMetPheSerAlaThrPhePro 428

Qy         1456 GAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT 1515
Db          429 AspGluIleGlnArgLeuAlaHisAspPheLeuArgGluAspPheLeuPheLeuThrVal 448

Qy         1516 GGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTC 1575
Db          449 GlyArgValGlyGlyAlaCysThrAspValThrGlnSerIleIleGlnValAspGlnAsp 468

Qy         1576 TCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT------GAAAGAACT 1629
Db          469 AspLysArgAlaLysLeuLeuGluLeuIleSerAspValAlaGluThrArgSerArgThr 488

Qy         1630 ATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAA 1689
Db          489 LeuValPheValGluThrLysArgGlyAlaAspPheLeuAlaCysMetLeuSerGlnGlu 508

Qy         1690 AAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTT 1749
Db          509 GlyCysProThrThrSerIleHisGlyAspArgLeuGlnGlnGluArgGluGlnAlaLeu 528

Qy         1750 GGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGG 1809
Db          529 ArgAspPheLysSerAlaValCysProIleLeuIleAlaThrSerValAlaAlaArgGly 548

Qy         1810 CTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAA 1869
Db          549 LeuAspIleProLysValGluHisValIleAsnTyrAspMetProLysGluIleAspGlu 568

Qy         1870 TATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTT 1929
Db          569 TyrValHisArgIleGlyArgThrGlyArgCysGlyAsnLeuGlyArgAlaThrThrPhe 588

Qy         1930 TTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCT 1989
Db          589 TyrAspAsnAsnLysAspGlyGluLeuAlaArgSerLeuValLysIleLeuSerGluAla 608

Qy         1990 CAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTC 2049
Db          609 GlnGlnGluValProGlyTrpLeuGluGluCysAlaGluSerAlaVal------------ 624

Qy         2050 AGTGGTAGTACAAGAGGAAACGTG------TTTGCATCAGTTGATACCAGAAAG------ 2097
Db          625 ---GlySerSerPheGlyLysGluGlyGlyPheGlyGlyArgAspPheArgLysArgGly 643

Qy         2098 -----------------------------------GGCAAGAGCACTTTGAACACAGCT 2121
Db          644 GlyArgAsnAspThrGlyPheGlyGlySerLysGlyGlyPheGlySerArgSerThrAla 663

Qy         2122 -----------------------GGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGAT 2157
Db          664 AspCysAspTyrAsnAspGlyGlyGlyPheGlyAsnThrAlaAlaValSerLysAspAsp 683

Qy         2158 GATGAGTCATGGGAT 2172
Db          684 AspAspSerTrpAsp 688

RESULT 13
O96069
ID   O96069      PRELIMINARY;      PRT;   659 AA.
AC   O96069;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   DEAD-box protein.
GN   CIDEAD1A.
OS   Ciona intestinalis.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=7719;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Takamura K., Fujimura M.;
RT   "Expression Patterns of Ascidian DEAD-Box Genes.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AB016604; BAA36711.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
```

```
DR   Pfam; PF00098; zf-CCHC; 3.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 3.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   659 AA;  71609 MW;  234F8DAC0D5D9699 CRC64;

Alignment Scores:
Pred. No.:              2.82e-110      Length:          659
Score:                  1522.00        Matches:         338
Percent Similarity:     61.54%         Conservative:    86
Best Local Similarity:  49.06%         Mismatches:      191
Query Match:            38.49%         Indels:          74
DB:                     5              Gaps:            17

US-09-714-865-15 (1-2172) x O96069 (1-659)

Qy    196 GGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG--------- 246
Db     15 GlyAsnThrAspPheArgGlyAlaSerSerPheAspAsnProTyrSerLysLysAspAsp 34

Qy    247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGA---AACAGAGGTTTTTCAAACAGCAGG--- 300
Db     35 GlyGlyPheGlySerGly------PheGlyAspSerArgGlyGlyAlaArgSerLysGly 52

Qy    301 ---TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGAT 357
Db     53 CysPheLysCysGlyGluGluGly------------HisMetSerArgGluCysProGln 68

Qy    358 AATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCA 417
Db     69 AsnThrGlySerGlyPheGlyAspSer---ArgGlyGlyAlaArg--------------- 82

Qy    418 GAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGA 477
Db     83 ---SerLysGlyCysPheLysCysGlyGluGluGlyHisMet------------------ 95

Qy    478 TTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGT 537
Db     96 --------------------------------SerArgGluCysProGlnAsnThrGly 104

Qy    538 GGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAA 597
Db    105 SerGlyPheGlyAspSerArg---------------------GlyGlyAlaArgSerLys 117

Qy    598 AGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATA 657
Db    118 GlyCysPheLysCysGlyGluGlu------GlyHisMetSerArgGluCysProGlnAsn 135

Qy    658 ACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACT 717
Db    136 ThrGlySerGlyGlyAspHisAlaAlaGlnGluTyrHisLysAlaGlyAspGlyAspLys 155

Qy    718 CAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTT 777
Db    156 ProArgProProLeu---TyrIleProProProProProGluAspGluValGluMetPhe 174

Qy    778 GCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
Db    175 AlaSerMetGlnArgGlyIleAsnPheGlyLysTyrAspAlaIleProValGluValSer 194

Qy    838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
Db    195 GlyLeuAsnAlaProLysCysIleSerThrPheGluMetAlaAsnLeuGlnGluThrIle 214

Qy    898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
Db    215 LeuValAsnValGlnLysAlaGlyTyrAspArgProThrProValGlnLysTyrSerIle 234

Qy    958 CCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACT 1017
Db    235 ProIleIleAsnAlaAspArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThr 254

Qy   1018 GCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGT 1077
Db    255 AlaAlaPheLeuLeuProValLeuThrLysLeuValGluSerGlyValLysSerSerGlu 274

Qy   1078 TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAAC 1137
Db    275 PheSerGluLysLysThrProGlnAlaIleIleIleGlyProThrArgGluLeuValAsn 294

Qy   1138 CAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATA 1197
Db    295 GlnIlePheLeuGluAlaArgLysPheSerArgSerThrIleIleHisProValValVal 314

Qy   1198 TATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTA 1257
Db    315 TyrGlyGlyThrSerValGlyTyrGlnIleArgAlaValGlnLysGlyCysAspValLeu 334

Qy   1258 TGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAG 1317
Db    335 IleAlaThrProGlyArgLeuMetAspPheIleAsnArgGlyLeuIleGlyLeuGluAsn 354

Qy   1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAA 1377
Db    355 ValArgPheIleIleLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluSerGlu 374

Qy   1378 ATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATG 1437
Db    375 IleArgLysLeuValSerLeuProGlyMetProGlnLysAsnAsnArgHisThrLeuMet 394

Qy   1438 TTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAAT 1497
Db    395 PheSerAlaThrPheProAspGluIleGlnLysLeuAlaHisAspPheLeuArgGluAsp 414

Qy   1498 TATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT 1557
Db    415 PheLeuPheLeuThrValGlyArgValGlyGlyAlaCysSerAspValThrGlnThrVal 434

Qy   1558 CTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG 1617
Db    435 IleSerValGluTrpLysAspLysArgSerLysLeuLeuGluLeuIleAlaAspValAsn 454

Qy   1618 GAT------GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCA 1671
Db    455 GluThrLysSerArgThrLeuValPheValGluThrLysArgGlyAlaAspPheLeuAla 474

Qy   1672 ACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGA 1731
Db    475 CysValLeuCysGlnGluAspPheProThrThrSerIleHisGlyAspArgLeuGlnGln 494

Qy   1732 GAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACT 1791
Db    495 AspArgGluGlnAlaLeuArgAspPheLysLeuAlaValCysProIleLeuValAlaThr 514

Qy   1792 TCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTT 1851
Db    515 SerValAlaAlaArgGlyLeuAspIleProLysValGluHisValIleAsnTyrAspMet 534

Qy   1852 CCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACT 1911
Db    535 ProArgGluIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnLeu 554

Qy   1912 GGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTA 1971
Db    555 GlyArgAlaThrThrPhePheAspAsnLysLysAspAlaAsnLeuAlaArgSerLeuVal 574

Qy   1972 AAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGT 2031
Db    575 LysIleLeuSerGluAlaGlnGlnGluValProGlyTrpLeuGlyGluCysAlaGluSer 594

Qy   2032 ACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACC 2091
Db    595 AlaValGlySerAsnPhe---GlyAlaGluLysGlyArg---PheGlyGlyLysAspLeu 612
```

```
Qy       2092 AGAAAG------------------------------------------GGCAAGAGCACT 2109
              |||:::                                          |||
Db        613 ArgGluArgGlyGlyGluGlyIleThrGlyPheGlySerMetGlyGlyGlySerGlyAla 632

Qy       2110 TTGAACACAGCT------------------------GGGTTTTCTTCTTCACGAGCTCCC 2145
                 :::||||||                        ||||||
Db        633 ArgSerThrAlaAspTyrAspTyrAsnAspGlyGlyGlyPheAspCysValGlyValAsn 652

Qy       2146 AATCCAGTAGATGATGAGTCATGGGAT 2172
              :::       ||||||:::||||||||
Db        653 Asp------AspAspAspSerTrpAsp 659
```

RESULT 14
O96068

```
ID   O96068      PRELIMINARY;      PRT;   669 AA.
AC   O96068;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   DEAD-box protein.
GN   CIDEAD1B.
OS   Ciona intestinalis.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=7719;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=20130953; PubMed=10664149;
RA   Fujimura M., Takamura K.;
RT   "Characterization of an ascidian DEAD-box gene, Ci-DEAD1: specific
RT   expression in the germ cells and its mRNA localization in the
RT   posterior-most blastomeres in early embryos.";
RL   Dev. Genes Evol. 210:64-72(2000).
CC   -!- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR   EMBL; AB016603; BAA36710.1; -.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF00098; zf-CCHC; 3.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 3.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   ATP-binding; Helicase; RNA-binding.
SQ   SEQUENCE   669 AA;  72785 MW;  1F5CA5BA546A16D8 CRC64;
```

```
Alignment Scores:
Pred. No.:               2.83e-110     Length:         669
Score:                   1522.00       Matches:        339
Percent Similarity:      61.32%        Conservative:   89
Best Local Similarity:   48.57%        Mismatches:     188
Query Match:             38.49%        Indels:         82
DB:                      5             Gaps:           18
```

US-09-714-865-15 (1-2172) x O96068 (1-669)

```
Qy        196 GGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG--------- 246
              ||||||   |||           :::      ||||||      :::
Db         15 GlyAsnThrAspPheArgGlyAlaSerSerPheAspAsnProTyrSerLysLysAspAsp 34

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGA---AACAGAGGTTTTTCAAACAGCAGG--- 300
              ||||||||||||   |||      ||||||   :::||||||   :::   |||:::
Db         35 GlyGlyPheGlySerGly------PheGlyAspSerArgGlyGlyAlaArgSerLysGly 52

Qy        301 ---TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGAT 357
                 |||:::   |||:::                        |||   :::|||
Db         53 CysPheLysCysGlyGluGluGly------------HisMetSerArgGluCysProGln 68

Qy        358 AATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCA 417
              |||         |||   |||   |||||||||   |||
Db         69 AsnThrGlySerGlyPheGlyAspSer---ArgGlyGlyAlaArg--------------- 82

Qy        418 GAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGA 477
                 :::   |||   ::::::   |||      |||
Db         83 ---SerLysGlyCysPheLysCysGlyGluGluGlyHisMet------------------ 95

Qy        478 TTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGT 537
                                                   ||||||   |||   ||||||
Db         96 ---------------------------------SerArgGluCysProGlnAsnThrGly 104

Qy        538 GGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAA 597
                 ||||||   |||                        |||   |||:::
Db        105 SerGlyPheGlyAspSerArg--------------------GlyGlyAlaArgSerLys 117

Qy        598 AGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATA 657
                       |||   |||      |||:::
Db        118 GlyCysPheLysCysGlyGluGlu------GlyHisMetSerArgGluCysProGlnAsn 135

Qy        658 ACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA------------ 705
              ||||||||||||            :::      :::||||||:::
Db        136 ThrGlySerGlyAspArgHisSerAsnAlaTyrPheLysGlyGlyAspHisAlaAlaGln 155

Qy        706 ------------AGTAGTGATACTCAAGGACCAAAA---GTGACCTACATACCCCCTCCT 750
                          :::   |||      |||:::      |||||||||||||||
Db        156 GluTyrHisLysAlaGlyAspGlyAspLysProArgProProLeuTyrIleProProPro 175

Qy        751 CCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAA 810
              |||||||||||||||      :::||||||      |||   ||||||||||||   |||
Db        176 ProProGluAspGluValGluMetPheAlaSerMetGlnArgGlyIleAsnPheGlyLys 195

Qy        811 TACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTT 870
              ||||||   |||   |||||||||||||||      :::||||||      |||   ||||||
Db        196 TyrAspAlaIleProValGluValSerGlyLeuAsnAlaProLysCysIleSerThrPhe 215

Qy        871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
              |||   |||||||||   :::|||:::      |||:::   ||||||||||||   :::
Db        216 GluMetAlaAsnLeuGlnGluThrIleLeuValAsnValGlnLysAlaGlyTyrAspArg 235

Qy        931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
              |||||||||||||||||||||||||||||||||   |||   |||||||||||||||
Db        236 ProThrProValGlnLysTyrSerIleProIleIleAsnAlaAspArgAspLeuMetAla 255

Qy        991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG 1050
              |||||||||||||||||||||||||||||||||||||||||||:::|||         :::
Db        256 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProValLeuThrLysLeu 275

Qy       1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATT 1110
              :::      |||:::   :::|||   |||   |||   :::   |||:::   ||||||
Db        276 ValGluSerGlyValLysSerSerGluPheSerGluLysLysThrProGlnAlaIleIle 295

Qy       1111 GTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTT 1170
              :::   ||||||||||||||||||||||||:::||||||||||||||||||||
Db        296 IleGlyProThrArgGluLeuValAsnGlnIlePheLeuGluAlaArgLysPheSerArg 315

Qy       1171 GGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGA 1230
                 |||   :::      ||||||:::||||||||||||      :::|||:::   ||||||
Db        316 SerThrIleIleHisProValValValTyrGlyGlyThrSerValGlyTyrGlnIleArg 335

Qy       1231 CAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATA 1290
                 :::   :::||||||::::::|||   ||||||||||||||||||||||   |||
Db        336 AlaValGlnLysGlyCysAspValLeuIleAlaThrProGlyArgLeuMetAspPheIle 355

Qy       1291 GGCAAAGAAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGC 1350
                 :::         |||||||||:::   ::::::::::::::|||||||||||||||||
Db        356 AsnArgGlyLeuIleGlyLeuGluAsnValArgPheIleIleLeuAspGluAlaAspArg 375

Qy       1351 ATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCA 1410
              ||||||||||||||||||      |||::::::||||||:::|||   ||||||||||||
Db        376 MetLeuAspMetGlyPheGluSerGluIleArgLysLeuValSerLeuProGlyMetPro 395

Qy       1411 TCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGG 1470
```

```
              |||       |||   |||||||||||||||||||||||||:::|||||||||:::
Db        396 GlnLysAsnAsnArgHisThrLeuMetPheSerAlaThrPheProAspGluIleGlnLys 415

Qy       1471 TTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGA 1530
              ||||||   :::||||||:::   ::::::||||||:::   ||||||:::|||||||||
Db        416 LeuAlaHisAspPheLeuArgGluAspPheLeuPheLeuThrValGlyArgValGlyGly 435

Qy       1531 GCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAG 1590
              ||||||   ||||||   ||||||||||:::   |||            ||||||   |||
Db        436 AlaCysSerAspValThrGlnThrValIleSerValGluTrpLysAspLysArgSerLys 455

Qy       1591 CTCGTTGAAATTCTGCGAAACATAGGGGAT------GAAAGAACTATGGTCTTTGTTGAA 1644
              |||:::|||::::::   ::::::   :::         ||||||:::||||||||||||
Db        456 LeuLeuGluLeuIleAlaAspValAsnGluThrLysSerArgThrLeuValPheValGlu 475

Qy       1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACA 1704
              ||||||:::   |||||||||   |||         |||||||||||         ||||||
Db        476 ThrLysArgGlyAlaAspPheLeuAlaCysValLeuCysGlnGluAspPheProThrThr 495

Qy       1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
              ||||||||||||||||||   |||::::::|||||||||||||||   ||||||:::
Db        496 SerIleHisGlyAspArgLeuGlnGlnAspArgGluGlnAlaLeuArgAspPheLysLeu 515

Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824
                       ||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        516 AlaValCysProIleLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleProLys 535

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884
              |||:::||||||||||||:::|||:::|||         |||||||||||||||||||||
Db        536 ValGluHisValIleAsnTyrAspMetProArgGluIleAspGluTyrValHisArgIle 555

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCG 1944
              ||||||||||||||||||||||||   |||||||||   :::|||||||||   :::
Db        556 GlyArgThrGlyArgCysGlyAsnLeuGlyArgAlaThrThrPhePheAspAsnLysLys 575

Qy       1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
              |||   :::||||||:::   |||||||||:::|||::::::|||||||||:::||||||
Db        576 AspAlaAsnLeuAlaArgSerLeuValLysIleLeuSerGluAlaGlnGlnGluValPro 595

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGA 2064
                 ||||||   |||   |||   |||                     |||   |||:::   :::
Db        596 GlyTrpLeuGlyGluCysAlaGluSerAlaValGlySerAsnPhe---GlyAlaGluLys 614

Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAG-------------------------- 2097
              |||      |||         |||   |||:::
Db        615 GlyArg---PheGlyGlyLysAspLeuArgGluArgGlyGlyGluGlyIleThrGlyPhe 633

Qy       2098 ---------------GGCAAGAGCACTTTGAACACAGCT-------------------- 2121
                             |||         :::||||||
Db        634 GlySerMetGlyGlyGlySerGlyAlaArgSerThrAlaAspTyrAspTyrAsnAspGly 653

Qy       2122 ---GGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
                 ||||||                     :::      ||||||:::|||||||||
Db        654 GlyGlyPheAspCysValGlyValAsnAsp------AspAspAspSerTrpAsp 669
```

RESULT 15
Q8QGD0

```
ID   Q8QGD0      PRELIMINARY;      PRT;   394 AA.
AC   Q8QGD0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   DEAD box RNA helicase Vasa.
OS   Pantodon buchholtzi (Butterflyfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC   Osteoglossiformes; Pantodontidae; Pantodon.
OX   NCBI_TaxID=8276;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21906633; PubMed=11909530;
RA   Knaut H., Steinbeisser H., Schwarz H., Nusslein-Volhard C.;
RT   "An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA
RT   Translation to the Germ Cells in the Zebrafish.";
RL   Curr. Biol. 12:454-466(2002).
DR   EMBL; AF479823; AAL87142.1; -.
KW   Helicase.
SQ   SEQUENCE   394 AA;  43662 MW;  78F9ADF1A4A0A324 CRC64;

Alignment Scores:
Pred. No.:               2.19e-108     Length:         394
Score:                   1497.00       Matches:        288
Percent Similarity:      84.75%        Conservative:   51
Best Local Similarity:   72.00%        Mismatches:     51
Query Match:             37.86%        Indels:         10
DB:                      13            Gaps:           3

US-09-714-865-15 (1-2172) x Q8QGD0 (1-394)

Qy        985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCT 1044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuGln 20

Qy       1045 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
                 :::|||         |||:::   :::||||||   |||:::||||||||||||
Db         21 GlnLeuMetThrSerGlyValAlaSerSerArgPheSerGluValGlnGluProGluAla 40

Qy       1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164
              |||:::|||||||||||||||||||:::||||||||||||||||||||||||||||||
Db         41 IleValValAlaProThrArgGluLeuIleAsnGlnIleTyrLeuGluAlaArgLysPhe 60

Qy       1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA 1224
              :::||||||||||||||||||   ||||||:::||||||||||||         ::::::
Db         61 AlaPheGlyThrCysValArgProValValValTyrGlyGlyThrSerThrMetTyrThr 80

Qy       1225 ATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGAT 1284
              ||||||:::::::::::::|||||||||||||||||   ||||||||||||||:::|||
Db         81 IleArgGluValLeuLysGlyCysAsnIleLeuCysGlyThrProGlyArgLeuLeuAsp 100

Qy       1285 ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCT 1344
              |||||||||:::   |||:::||||||   :::::::::||||||||||||||||||||
Db        101 IleIleGlyArgGlyLysValGlyLeuAsnLysLeuArgTyrLeuValLeuAspGluAla 120

Qy       1345 GATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
              ||||||||||||||||||||||||   |||:::|||::::::|||::::::   ||||||
Db        121 AspArgMetLeuAspMetGlyPheGluProAspMetArgArgLeuValAlaSerProGly 140

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
              ||||||   ||||||:::||||||||||||||||||||||||:::||||||:::|||
Db        141 MetProProLysGluGluArgGlnThrLeuMetPheSerAlaThrTyrProGluAspIle 160

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
              |||||||||||||||:::|||||||||   :::|||||||||:::   ||||||||||||
Db        161 GlnArgLeuAlaAlaAspPheLeuLysArgAspTyrLeuPheLeuValValGlyGlnVal 180

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
              ||||||||||||   ||||||:::|||   |||::::::|||   |||:::|||||||||
Db        181 GlyGlyAlaCysAsnAspValGluGlnHisValValGluValSerGlnTyrSerLysArg 200

Qy       1585 GAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAA 1644
              |||:::||||||||||||||||||         |||:::||||||||||||||||||||||
Db        201 GluGlnLeuValGluIleLeuArgThrThrGlyAsnGluArgThrMetValPheValGlu 220

Qy       1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACA 1704
              ||||||:::||||||||||||   ||||||||||||||||||||:::||||||||||||
Db        221 ThrLysArgLysAlaAspPheIleAlaThrPheLeuCysGlnGluGlnIleSerThrThr 240

Qy       1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SerIleHisGlyAspArgGluGlnArgGluArgGluGlnAlaLeuGlyAspPheArgSer 260

Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        261 GlyLysCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluHis 280

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884
```

```
               ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db         281 ValGlnHisValValAsnPheAspLeuProSerThrIleAspGluTyrValHisArgIle 300

Qy        1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCG 1944
               |||||||||||||||||||||||||||||:::||||||||||||||||||   |||:::
Db         301 GlyArgThrGlyArgCysGlyAsnThrGlyLysAlaIleSerPhePheAspProGluAla 320

Qy        1945 GATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
               |||      ||||||:::   ||||||||:::|||:::   |||||||||:::||||||
Db         321 AspAlaProLeuAlaArgSerLeuValLysIleLeuSerGlyAlaGlnGlnGluValPro 340

Qy        2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACA--- 2061
                  ||||||||||||||||||||||   :::                |||:::|||
Db         341 LysTrpLeuGluGluIleAlaPheSerAlaHis---------------GlyThrThrAla 355

Qy        2062 ---------AGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTG 2112
                        ||||||   ||||||||||||   ||||||||||||   :::
Db         356 PheAsnProArgGlyLysValPheAlaSerThrAspThrArgLys---GluGlyAlaPhe 374

Qy        2113 AACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
                     |||                           ||||||     ||||||   |||:::
Db         375 ProArgAlaGluProValGlnProAlaValGlnAsnProAspAlaAspGluGluTrpGlu 394
```

Search completed: June 10, 2003, 16:42:45
Job time : 150 secs

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:         June 10, 2003, 16:32:00 ; Search time 47.5 Seconds
                                          (without alignments)
                                          8791.746 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  3954
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters:      566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111758_26034/app_query.fasta_1.2311
-DB=PIR_73 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_55_@runat_05062003_111758_26034 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      PIR_73:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3323	84.0	635	2	T46407	probable RNA helic
2	3247.5	82.1	713	2	JC2534	RVLG protein - rat
3	2857.5	72.3	637	2	I49638	probable RNA helic
4	1959.5	49.6	700	2	I51235	DEAD box protein -
5	1274	32.2	661	2	A58768	ATP-dependent RNA
6	1229	31.1	662	1	I84741	RNA helicase - mou
7	1224.5	31.0	697	1	S13654	ATP-dependent RNA
8	1197	30.3	660	2	A32378	probable ATP-depen
9	1176	29.7	688	2	T48796	probable ATP-depen
10	1126.5	28.5	604	2	S13653	ATP-dependent RNA
11	1109	28.0	636	2	T43543	probable ATP-depen
12	1108	28.0	617	2	S62003	probable ATP-depen
13	1100	27.8	633	2	H84854	probable ATP-depen
14	1095.5	27.7	646	2	T45677	ATP-dependent RNA
15	1083.5	27.4	707	2	A48686	probable RNA helic
16	1069.5	27.0	603	2	T45671	ATP-dependent RNA
17	1026.5	26.0	604	2	C87818	protein glh-1 [imp
18	1026.5	26.0	604	2	T15132	ATP-dependent RNA
19	992.5	25.1	720	2	T15231	germline RNA helic
20	820.5	20.8	550	1	S14048	RNA helicase dbp2
21	815	20.6	546	1	S13757	RNA helicase DBP2
22	804	20.3	561	2	T22917	probable ATP-depen
23	797.5	20.2	619	2	T52137	ATP-dependent DEAD
24	796	20.1	575	1	S11485	RNA helicase - fru
25	790	20.0	501	2	B96593	probable ethylene-
26	777.5	19.7	614	1	I48385	RNA helicase TNZ2
27	777.5	19.7	614	2	JC1087	RNA helicase, ATP-
28	771.5	19.5	1156	2	T43326	germline RNA helic
29	771.5	19.5	1172	2	T32759	hypothetical prote
30	754	19.1	607	1	S42639	ATP-dependent RNA
31	754	19.1	650	2	S72367	ATP-dependent RNA
32	753	19.0	713	2	T48634	DRH1 DEAD box prot
33	741.5	18.8	1032	2	A57514	RNA helicase HEL11
34	713	18.0	748	2	H84913	probable ATP-depen
35	713	18.0	1166	2	H86341	hypothetical prote
36	712	18.0	446	2	E82835	ATP-dependent RNA
37	699	17.7	733	2	H84748	hypothetical prote
38	696.5	17.6	566	2	S53813	RNA helicase - sli
39	693.5	17.5	523	1	S30805	probable RNA helic
40	692.5	17.5	578	2	S67386	probable ATP-depen
41	688.5	17.4	537	2	A86444	probable RNA helic
42	685	17.3	454	2	E64816	probable ATP-depen
43	682.5	17.3	701	2	T15942	hypothetical prote
44	682	17.2	455	2	AI0599	probable ATP-depen
45	681.5	17.2	455	2	C90738	probable ATP-depen

ALIGNMENTS

```
RESULT 1
T46407
probable RNA helicase protein DKFZp434B1122.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
C;Accession: T46407
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23034
A;Accession: T46407
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-635 <AAA>
A;Cross-references: EMBL:AL137462; PIDN:CAB70750.1
A;Experimental source: adult testis; clone DKFZp434B1122
C;Genetics:
A;Note: DKFZp434B1122.1

Alignment Scores:
Pred. No.:              2.1e-230        Length:         635
Score:                  3323.00         Matches:        633
Percent Similarity:     99.84%          Conservative:   1
Best Local Similarity:  99.69%          Mismatches:     1
Query Match:            84.04%          Indels:         0
DB:                     2               Gaps:           0

US-09-714-865-15 (1-2172) x T46407 (1-635)

Qy        268 AGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTC 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerPheGlyAsnArgGlyPheSerAsnSerArgPheGluAspGlyAspSerSerGlyPhe 20

Qy        328 TGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAG 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 TrpArgGluSerSerAsnAspCysGluAspAsnProThrArgAsnArgGlyPheSerLys 40

Qy        388 AGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGA 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         41 ArgGlyGlyTyrArgAspGlyAsnAsnSerGluAlaSerGlyProTyrArgArgGlyGly 60

Qy        448 AGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTA 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ArgGlySerPheArgGlyCysArgGlyGlyPheGlyLeuGlySerProAsnAsnAspLeu 80

Qy        508 GACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTA 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AspProAspGluCysMetGlnArgThrGlyGlyLeuPheGlySerArgArgProValLeu 100

Qy        568 AGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGT 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 SerGlyThrGlyAsnGlyAspThrSerGlnSerArgSerGlySerGlySerGluArgGly 120

Qy        628 GGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlyTyrLysGlyLeuAsnGluGluValIleThrGlySerGlyLysAsnSerTrpLysSer 140

Qy        688 GAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCT 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GluAlaGluGlyGlyGluSerSerAspThrGlnGlyProLysValThrTyrIleProPro 160

Qy        748 CCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGAC 807
              |||·||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ProProProGluAspGluAspSerIlePheAlaHisTyrGlnThrGlyIleAsnPheAsp 180

Qy        808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LysTyrAspThrIleLeuValGluValSerGlyHisAspAlaProProAlaIleLeuThr 200

Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 PheGluGluAlaAsnLeuCysGlnThrLeuAsnAsnAsnIleAlaLysAlaGlyTyrThr 220

Qy        928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 LysLeuThrProValGlnLysTyrSerIleProIleIleLeuAlaGlyArgAspLeuMet 240

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuAlaHis 260

Qy       1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 MetMetHisAspGlyIleThrAlaSerArgPheLysGluLeuGlnGluProGluCysIle 280

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 IleValAlaProThrArgGluLeuValAsnGlnIleTyrLeuGluAlaArgLysPheSer 300

Qy       1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PheGlyThrCysValArgAlaValValIleTyrGlyGlyThrGlnLeuGlyHisSerIle 320

Qy       1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ArgGlnIleValGlnGlyCysAsnIleLeuCysAlaThrProGlyArgLeuMetAspIle 340

Qy       1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 IleGlyLysGluLysIleGlyLeuLysGlnIleLysTyrLeuValLeuAspGluAlaAsp 360

Qy       1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ArgMetLeuAspMetGlyPheGlyProGluMetLysLysLeuIleSerCysProGlyMet 380

Qy       1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ProSerLysGluGlnArgGlnThrLeuMetPheSerAlaThrPheProGluGluIleGln 400

Qy       1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGT 1527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 ArgLeuAlaAlaGluPheLeuLysSerAsnTyrLeuPheValAlaValGlyGlnValGly 420

Qy       1528 GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAA 1587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GlyAlaCysArgAspValGlnGlnThrValLeuGlnValGlyGlnPheSerLysArgGlu 440

Qy       1588 AAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACT 1647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LysLeuValGluIleLeuArgAsnIleGlyAspGluArgThrMetValPheValGluThr 460

Qy       1648 AAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGT 1707
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        461 LysLysLysAlaAspPheIleAlaThrPheLeuCysGlnGluLysIleSerThrThrSer 480

Qy       1708 ATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGA 1767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 IleHisGlyAspArgGluGlnArgGluArgGluGlnAlaLeuGlyAspPheArgPheGly 500

Qy       1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LysCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluAsnVal 520

Qy       1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GlnHisValIleAsnPheAspLeuProSerThrIleAspGluTyrValHisArgIleGly 540

Qy       1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGluSerAsp 560

Qy       1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 AsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValProAla 580

Qy       2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 TrpLeuGluGluIleAlaPheSerThrTyrIleProGlyPheSerGlySerThrArgGly 600

Qy       2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AsnValPheAlaSerValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGlyPhe 620

Qy       2128 TCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
              ||||||||:::||||||||||||||||||||||||||||||||||
Db        621 SerSerSerGlnAlaProAsnProValAspAspGluSerTrpAsp 635
```

RESULT 2
JC2534
RVLG protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-May-1995 #sequence_revision 14-Jul-1995 #text_change 02-Feb-2001
C;Accession: JC2534
R;Komiya, T.; Tanigawa, Y.
Biochem. Biophys. Res. Commun. 207, 405-410, 1995
A;Title: Cloning of a gene of the DEAD Box protein family which is specifically expressec
A;Reference number: JC2534; MUID:95160706; PMID:7857296
A;Accession: JC2534
A;Molecule type: mRNA
A;Residues: 1-713 <KOM>
A;Cross-references: GB:S75275; NID:g806463; PIDN:AAB33364.1; PID:g806464
C;Comment: This protein contains a conserved DEAD box.
C;Keywords: ATP; nucleotide binding; P-loop
F;317-324/Region: nucleotide-binding motif A (P-loop)
F;427-432/Region: nucleotide-binding motif B
F;431-434/Region: DEAD motif

```
Alignment Scores:
Pred. No.:               5.65e-225     Length:        713
Score:                   3247.50       Matches:       625
Percent Similarity:      90.55%        Conservative:  36
Best Local Similarity:   85.62%        Mismatches:    46
Query Match:             82.13%        Indels:        23
DB:                      2             Gaps:          6
```

```
US-09-714-865-15 (1-2172) x JC2534 (1-713)

Qy          1 ATGGGGGATGAAGATTGGGAAGCAGAAATC---AACCCTCATATGTCTTCCTATGTTCCC 57
              |||||||||||||||||||||||||||||||   |||||:::|||||||||||||||
Db          1 MetGlyAspGluAspTrpGluAlaGluIleLeuLysProHisValSerSerTyrValPro 20

Qy         58 ATATTTGAGAAGGATAGGTAT---TCTGGAGAAAATGGAGACAATTTTAACAGGACTCCA 114
              :::|||||||||||:::|||   ||||||   |||||||||   ||||||||||||
Db         21 ValPheGluLysAspLysTyrSerSerGlyAlaAsnGlyAspThrPheAsnArgThrSer 40

Qy        115 GCTTCATCATCAGAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGA 174
              ||||||||||||||||||:::|||||||||||  ||||||||||||||:::||||||
Db         41 AlaSerSerSerGluMetGluAspGlyProSerGlyArgAspHisPheMetArgSerGly 60

Qy        175 TTTGCCTCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT 234
              |||:::||||||||||||   |||||||||||||   ||||||   :::||||||:::
Db         61 PheSerSerGlyArgAsnLeuGlyAsnArgAspIleGlyGluSerSerLysArgGluThr 80

Qy        235 ACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAAC 294
              |||||||||   ||||||||||||   ||||||   ||||||||||||||||||   |||
Db         81 ThrSerThrThrGlyGlyPheGlyArgGlyLysGlyPheGlyAsnArgGlyPheLeuAsn 100

Qy        295 AGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAA 354
              ::::::||||||:::||||||||||||||||||||||:::||||||:::||||||||||||
Db        101 AsnLysPheGluGluGlyAspSerSerGlyPheTrpLysGluSerThrAsnAspCysGlu 120

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
              |||   ||||||:::||||||||||||||||||||||||||||   |||||||||:::
Db        121 AspThrGlnThrArgSerArgGlyPheSerLysArgGlyGlyTyrProAspGlyAsnAsp 140

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGA 474
              |||||||||||||||||:::|||||||||||||||||   |||
Db        141 SerGluAlaSerGlyProPheArgArgGlyGlyArgAspSer------------------ 154

Qy        475 GGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACT 534
                                              :::   |||   |||:::   ||||||
Db        155 ---------------------------GluTyrAspGlnAspGlnGlySerGlnArgGly 165

Qy        535 GGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCT 594
              |||||||||||||||||||||:::|||   |||   :::|||:::|||||||||
Db        166 GlyGlyLeuPheGlySerArgLysProAlaAlaSerAspSerGlySerGlyAspThrPhe 185

Qy        595 CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654
              |||||||||||||||:::         ||||||   ||||||||||||||||||||||||
Db        186 GlnSerArgSerGlyAsnAla------ArgGlyAlaTyrLysGlyLeuAsnGluGluVal 203

Qy        655 ATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGAT 714
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValThrGlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyGluSerSerAsp 223

Qy        715 ACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATC 774
                    |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 IleGlnGlyProLysValThrTyrIleProProProProProGluAspGluAspSerIle 243

Qy        775 TTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 PheAlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluVal 263

Qy        835 TCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACA 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 SerGlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThr 283

Qy        895 CTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGT 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 LeuAsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSer 303

Qy        955 ATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAG 1014
              ||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 IleProIleValLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLys 323

Qy       1015 ACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGT 1074
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        324 ThrAlaAlaPheLeuLeuProIleLeuAlaHisMetMetArgAspGlyIleThrAlaSer 343

Qy       1075 CGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTC 1134
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        344 ArgPheLysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuIle 363

Qy       1135 AACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTT 1194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 AsnGlnIleTyrLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaValVal 383

Qy       1195 ATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATA 1254
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        384 IleTyrGlyGlyThrGlnPheGlyHisSerIleArgGlnIleValGlnGlyCysAsnIle 403

Qy       1255 TTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAA 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 LeuCysAlaThrProGlyArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLys 423

Qy       1315 CAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCA 1374
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GlnValLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlyPro 443

Qy       1375 GAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTT 1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GluMetLysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeu 463

Qy       1435 ATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCA 1494
              :::|||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        464 LeuPheSerAlaThrPheProGluGluIleGlnArgLeuAlaGlyGluPheLeuLysSer 483

Qy       1495 AATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACC 1554
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        484 AsnTyrLeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnSer 503

Qy       1555 GTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATA 1614
              :::||||||||||||   |||:::   ||||||||||||||||||||||||||||||
Db        504 IleLeuGlnValGlyProValPheLysLysArgLysLeuValGluIleLeuArgAsnIle 523

Qy       1615 GGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
              ||||||||||||   ||||||||||||||||||||||||||||||||||||   ||||||
Db        524 GlyAspGluArgProMetValPheValGluThrLysLysLysAlaAspPheIleAlaThr 543

Qy       1675 TTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 PheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGlu 563

Qy       1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        564 ArgGluGlnAlaLeuGlyAspPheArgCysGlyLysCysProValLeuValAlaThrSer 583

Qy       1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        584 ValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAsnLeuPro 603

Qy       1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 SerThrIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGly 623

Qy       1915 AGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAA 1974
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        624 ArgAlaIleSerPhePheAspThrGluSerAspAsnHisLeuAlaGlnProLeuValLys 643

Qy       1975 GTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACA 2034
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        644 ValLeuSerAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerSer 663

Qy       2035 TACATT---CCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACC 2091
              |||      |||   ||||||   ||||||||||||   ||||||||||||   ||||||
Db        664 TyrAlaProProSerPheSerAsnSerThrArgGlyAlaValPheAlaSerPheAspThr 683

Qy       2092 AGAAAG---------GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
              ||||||         ||||||:::||||||||||||||||||   ||||||::::::|||
Db        684 ArgLysAsnPheGlnGlyLysAsnThrLeuAsnThrAlaGlyIleSerSerAlaGlnAla 703
```

```
Qy       2143 CCCAATCCAGTAGATGATGAGTCATGGGAT 2172
              ||||||||||||||||||||||||||||||
Db        704 ProAsnProValAspAspGluSerTrpAsp 713

RESULT 3
I49638
probable RNA helicase protein - mouse (fragment)
N;Alternate names: Drosophila vasa homolog
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Feb-2001
C;Accession: I49638
R;Fujiwara, Y.; Komiya, T.; Kawabata, H.; Sato, M.; Fujimoto, H.; Furusawa, M.; Noce, T.
Proc. Natl. Acad. Sci. U.S.A. 91, 12258-12262, 1994
A;Title: Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosop
A;Reference number: I49638; MUID:95083681; PMID:7991615
A;Accession: I49638
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-637 <RES>
A;Cross-references: GB:D14859; NID:g286074; PIDN:BAA03584:1; PID:g286075
C;Keywords: ATP; nucleotide binding; P-loop
F;246-253/Region: nucleotide-binding motif A (P-loop)
F;356-361/Region: nucleotide-binding motif B
F;360-363/Region: DEAD motif

Alignment Scores:
Pred. No.:              5.25e-197     Length:        637
Score:                  2857.50       Matches:       549
Percent Similarity:     91.39%        Conservative:  35
Best Local Similarity:  85.92%        Mismatches:    48
Query Match:            72.27%        Indels:        8
DB:                     2             Gaps:          4

US-09-714-865-15 (1-2172) x I49638 (1-637)

Qy        253 TTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGT 312
              |:||||   |||||   ||||||||||||||||   |||::::::||||||:::|||
Db          1 PheGlyArgGlyLysGlyPheGlyAsnArgGlyPheLeuAsnAsnLysPheGluGluGly 20

Qy        313 GATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAAC 372
              ||||||||||||||||||:::||||||:::||||||||||||||||||   ||||||:::
Db         21 AspSerSerGlyPheTrpLysGluSerAsnAsnAspCysGluAspAsnGlnThrArgSer 40

Qy        373 AGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCA 432
              ||||||||||||||||||||||||   :::||||||||||:::||||||||||||||||||
Db         41 ArgGlyPheSerLysArgGlyGlyCysGlnAspGlyAsnAspSerGluAlaSerGlyPro 60

Qy        433 TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGT 492
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PheArgArgGlyGlyArgGlySerPheArgGlyCysArgGlyGlyPheGlyLeuGlyArg 80

Qy        493 CCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCT 552
              ||||||::::::   |||   |||:::         |||      ||||||
Db         81 ProAsnSerGluSerAspGlnAspGlnGlyThrGlnCysGlyGlyGlyPheLeuValLeu 100

Qy        553 AGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGT 612
                 :::|||   |||   :::|||||||||||||||   ||||||||||||||||||
Db        101 GlyLysProAlaAlaSerAspSerGlyAsnGlyAspThrTyrGlnSerArgSerGlySer 120

Qy        613 GGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAG 672
              |||                  |||||||||||||||||||||||||:::||||||||||||
Db        121 Gly------ArgGlyGlyTyrLysGlyLeuAsnGluGluValValThrGlySerGlyLys 138

Qy        673 AATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTG 732
              |||||||||||||||||   ||||||||||||||||||||:::|||||||||||||||
Db        139 AsnSerTrpLysSerGluThrGluGlyGlyGluSerSerAspSerGlnGlyProLysVal 158

Qy        733 ACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACA 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ThrTyrIleProProProProProGluAspGluAspSerIlePheAlaHisTyrGlnThr 178

Qy        793 GGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCA 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 GlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSerGlyHisAspAlaPro 198

Qy        853 CCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCT 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 ProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeuAsnAsnAsnIleArg 218

Qy        913 AAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCA 972
              ||||||||||||||||||||||||||||||||||||:::|||||||||:::||||||
Db        219 LysAlaGlyTyrThrLysLeuThrProValGlnLysTyrThrIleProIleValLeuAla 238

Qy        973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 GlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeu 258

Qy       1033 CCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAG 1092
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        259 ProIleLeuAlaHisMetMetArgAspGlyIleThrAlaSerArgPheLysGluLeuGln 278

Qy       1093 GAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAA 1152
              ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        279 GluProGluCysIleIleValAlaProThrArgGluLeuIleAsnGlnIleTyrLeuGlu 298

Qy       1153 GCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAG 1212
              ||||||||||||||||||||||||||||   :::||||||||||||||||||||||||||
Db        299 AlaArgLysPheSerPheGlyThrCysValIleSerValValIleTyrGlyGlyThrGln 318

Qy       1213 CTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGA 1272
                 ||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        319 PheGlyHisSerValArgGlnIleValGlnGlyCysAsnIleLeuCysAlaThrProGly 338

Qy       1273 AGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTT 1332
              ||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        339 ArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLysGlnValLysTyrLeuVal 358

Qy       1333 TTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATT 1392
              ||||||||||||||   ||||||||||||||||||   ||||||:::||||||||||||
Db        359 LeuAspGluAlaAspSerMetLeuAspMetGlyPheAlaProGluIleLysLysLeuIle 378

Qy       1393 TCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTT 1452
              ||||||||||||||||||||||||||||   |||||||||:::||||||||||||||||
Db        379 SerCysProGlyMetProSerLysGluGlnHisGlnThrLeuLeuPheSerAlaThrPhe 398

Qy       1453 CCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCT 1512
              ||||||||||||||||||||||||   :::|||||||||||||||||||||||||||||
Db        399 ProGluGluIleGlnArgLeuAlaGlyAspPheLeuLysSerAsnTyrLeuPheValAla 418

Qy       1513 GTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAG 1572
              |||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db        419 ValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThrIleLeuGlnValGlyGln 438

Qy       1573 TTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATG 1632
              :::   |||   :::   |||:::         |||||||||||||||||||||||
Db        439 TyrGlnLysGluLysSerLeuLeuArgPheTyrGluAsnIleGlyAspGluArgThrMet 458

Qy       1633 GTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAA 1692
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db        459 ValPheValGluThrLysLysLysAlaAspPheIleAlaThrPheLeuCysGlnGluLys 478

Qy       1693 ATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGA 1752
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 IleSerSerThrSerIleHisGlyAspArgGluGlnArgGluArgGluGlnAlaLeuGly 498

Qy       1753 GATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTG 1812
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 AspPheArgCysGlyLysCysProValLeuValAlaThrSerValAlaAlaArgGlyLeu 518

Qy       1813 GATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATAT 1872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 AspIleGluAsnValGlnHisValIleAsnPheAspLeuProSerThrIleAspGluTyr 538

Qy       1873 GTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTT 1932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 ValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePhe 558
```

```
Qy       1933 GATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAA 1992
              |||    :::|||||||||||||||||||||||||||||||||||||:::|||||||||
Db        559 AspThrAspSerAspAsnHisLeuAlaGlnProLeuValLysValLeuSerAspAlaGln 578

Qy       1993 CAGGATGTTCCTGCATGGTTGGAGAAATTGCCTTTAGTACATACATT---CCTGGCTTC 2049
              ||||||||||||||||||||||||||||||||||||||||||:::    |||    |||
Db        579 GlnAspValProAlaTrpLeuGluGluIleAlaPheSerThrTyrValProProSerPhe 598

Qy       2050 AGTGGTAGTACAAGAGGAAAC---GTGTTTGCATCAGTTGATACCAGAAA---------- 2096
              |||.  |||||||||||       |||||||||||||||||||||||
Db        599 SerSerSerThrArgGlyGlyAlaValPheAlaSerValAspThrArgLysAsnTyrGln 618

Qy       2097 GGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGT 2153
              |||:::  |||  |||:::|||      ||||||||||||||||||||||||||
Db        619 GlyLysAlaHisValGluTyrSerGlyAspPhePhePheThrSerSerGlnSerSer 637
```

RESULT 4
I51235
DEAD box protein - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Feb-2001
C;Accession: I51235
R;Komiya, T.; Itoh, K.; Ikenishi, K.; Furusawa, M.
Dev. Biol. 162, 354-363, 1994
A;Title: Isolation and characterization of a novel gene of the DEAD box protein family w
A;Reference number: I51235; MUID:94200507; PMID:8150200
A;Accession: I51235
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-700 <KOM>
A;Cross-references: GB:S69534; NID:g2896106; PIDN:AAC03114.1; PID:g2896107
C;Genetics:
A;Gene: XVLG1
C;Keywords: ATP; nucleotide binding; P-loop
F;318-325/Region: nucleotide-binding motif A (P-loop)
F;428-433/Region: nucleotide-binding motif B
F;432-435/Region: DEAD motif

Alignment Scores:

Pred. No.:	1.44e-132	Length:	700
Score:	1959.50	Matches:	409
Percent Similarity:	68.17%	Conservative:	105
Best Local Similarity:	54.24%	Mismatches:	153
Query Match:	49.56%	Indels:	87
DB:	2	Gaps:	18

US-09-714-865-15 (1-2172) x I51235 (1-700)

```
Qy          7 GATGAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG 66
              :::|||:::|||:::   |||||            :::|||||||||   ||
Db          2 GluGluAsnTrpAspThrGluIleGluThrGluLysProThrTyrValProAsnPhe--- 20

Qy         67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
                          ||||||   ||||||:::       :::      |||::::::
Db         21 ---SerThrLeuGluThrGluAsnThrAspAsnTyr------SerAlaTyrSerAsnAsn 37

Qy        127 GAAATGGATGATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGG 186
              ::::::::::::                                  :::    |||
Db         38 AspIleAsnAsn------------------------------GlnAsnTyrAspSerGlu 47

Qy        187 CGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACAATG 246
              |||:::||||||||||||      |||      ::::::|||            |||
Db         48 ArgSerPheGlyAsnArg---GlyGlyTyrArgSerGluArgSerArgProSerAsnPhe 66

Qy        247 GGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAA 306
                                          |||||||||         ||||||   |||
Db         67 ------------------------------AsnArgGly---------SerArgThrGlu 73

Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------------- 351
                 |||      |||||              :::::::::
Db         74 ArgGlyArgGlyArgGlyPheGlyThrAsnArgAsnAspAsnTyrSerSerGluArgAsp 93

Qy        352 ---------GAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
                       :::|||       :::   |||||||||      ||||||||||||
Db         94 ValPheGlyAspAspGluArgAspGlnArgArgGlyPheProGlyArgGlyGlyTyrAsn 113

Qy        403 GATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGA 462
                 :::   :::            :::|||       |||||||||   ||||||
Db        114 GlyAsnGluAspGlyGlnLysProAsnAlaPheArg------GlyArgGlyGlyPheArg 131

Qy        463 ---------------------------GGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCA 495
                                         |||   ||||||||||||        :::
Db        132 AsnGluAsnGluGlnArgArgGlyPheGlyGluArgGlyGlyPheArgSerGluAsnGly 151

Qy        496 AATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTGGTTCTAGA 555
                    :::   |||                                 |||   ||||||:::
Db        152 GlnArgAsnPheAsp--------------------AsnArgGlyAspPheGlyAsnSer 164

Qy        556 ---------------AGACCAGTATTAAGTGGCACAGGT---------AATGGTGATACT 591
                             ||||||       |||   |||          |||   ||||||
Db        165 GlyGluGluGluAspArgProArgSerTyrGlyArgGlyGlyPheAsnAsnSerAspThr 184

Qy        592 TCT------------CAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGT 639
                 :::        |||   ||||||:::   ||||||||||||||||
Db        185 GlyGlyArgGlyArgArgGlyGlyArgGlyGlyGlySerGlnTyrGlyGlyTyrLysGly 204

Qy        640 TTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGA 699
                  ||||||||||||             |||      |||   ||||||:::   ||||||
Db        205 ArgAsnGluGluVal-----------GlyValGluSerGlyLysSerGlnGluGluGly 220

Qy        700 GGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAG 759
                  |||      |||   :::   ||||||||||||||||||||||||||||||:::
Db        221 AsnGlu---LysAspGluLysProLysLysValThrTyrIleProProProProProAsp 239

Qy        760 GATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACT 819
                 ||||||:::||||||      ||||||:::||||||||||||||||||||||||
Db        240 GlyGluAspAsnIlePheArgGlnTyrGlnSerGlyIleAsnPheAspLysTyrAspGlu 259

Qy        820 ATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCT 879
              |||||||||:::|||:::|||   |||   |||||||||||||||||||||||||||||
Db        260 IleLeuValAspValThrGlyLysAspValProProAlaIleLeuThrPheGluGluAla 279

Qy        880 AATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCT 939
              |||||||||:::||||||      |||:::|||:::|||||||||   ||||||||||||
Db        280 AsnLeuCysGluThrLeuArgArgAsnValAlaArgAlaGlyTyrValLysLeuThrPro 299

Qy        940 GTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAA 999
              |||||||||:::||||||||||||||||:::||||||||||||||||||||||||||||
Db        300 ValGlnLysHisSerIleProIleIleMetAlaGlyArgAspLeuMetAlaCysAlaGln 319

Qy       1000 ACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGAT 1059
              ||||||||||||||||||||||||||||||||||||||||||:::::|||||||:::::
Db        320 ThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerTyrMetMetAsnGlu 339

Qy       1060 GGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCA 1119
              |||||||||||||||||::::::   :::|||||||||||||||     ||||||:::||||||
Db        340 GlyIleThrAlaSerGlnTyrLeuGlnLeuGlnGluProGluAlaIleIleIleAlaPro 359

Qy       1120 ACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGT 1179
              ||||||||||||:::||||||||||||||||||:::||||||||||||||||:::||||||||
Db        360 ThrArgGluLeuIleAsnGlnIleTyrLeuAspAlaArgLysPheSerTyrGlyThrCys 379

Qy       1180 GTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTA 1239
              ||||||   ||||||:::|||||||||   |||         |||::::::|||    :::
Db        380 ValArgProValValValTyrGlyGlyIleGlnProValHisAlaMetArgAspValGlu 399

Qy       1240 CAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAA 1299
              :::||||||||||||||||||||||||||||||||||||:::||||||:::   ||||||
Db        400 LysGlyCysAsnIleLeuCysAlaThrProGlyArgLeuLeuAspIleValSerLysGlu 419

Qy       1300 AAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGAT 1359
              ||||||||||||    :::::::::||||||||||||||||||||||||||||||||||||
Db        420 LysIleGlyLeuSerLysLeuArgTyrLeuValLeuAspGluAlaAspArgMetLeuAsp 439

Qy       1360 ATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAA 1419
```

```
                |||||||||   ||||||::::::||||||::::::   ||||||||||||:::||||||
Db         440 MetGlyPheAlaProGluIleGluLysLeuMetThrLysProGlyMetProThrLysGlu 459

Qy        1420 CAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCA 1479
               :::||||||||||||||||||||||||||||:::||||||||||||:::|||||||||:::
Db         460 LysArgGlnThrLeuMetPheSerAlaThrTyrProGluGluIleArgArgLeuAlaSer 479

Qy        1480 GAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGA 1539
                  :::|||||||||   :::||||||||   ||||||   |||||||||||||||
Db         480 AsnTyrLeuLysSerGluHisLeuPheValValValGlyLeuValGlyGlyAlaCysSer 499

Qy        1540 GATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAA 1599
               ||||||   ||||||||||||::::::   :::         |||   |||||||||:::|||
Db         500 AspValAlaGlnThrValLeuGluMetArgGluAsnGlyLysMetGluLysLeuLeuGlu 519

Qy        1600 ATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCA 1659
               ||||||::::::            ||||||||||||:::||||||   |||||||||||||||
Db         520 IleLeuLysSerSerGluLysGluArgThrMetIlePheValAsnThrLysLysLysAla 539

Qy        1660 GATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGAT 1719
               ||||||   |||   :::|||||||||||||||   |||:::||||||||||||||||||
Db         540 AspPheIleAlaGlyTyrLeuCysGlnGluLysPheSerSerThrSerIleHisGlyAsp 559

Qy        1720 CGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTT 1779
               |||||||||   :::||||||   ||||||   |||||||||   |||||||||   |||
Db         560 ArgGluGlnTyrGlnArgGluSerAlaLeuTrpAspPheArgThrGlyLysCysThrVal 579

Qy        1780 CTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATC 1839
               :::|||   |||:::||||||||||||||||||||||||||||||||||||||||||||||
Db         580 IleValCysThrAlaValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIle 599

Qy        1840 AATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
               |||:::|||:::|||         :::||||||||||||||||||||||||||||||||||
Db         600 AsnTyrAspValProLysGluValAspGluTyrValHisArgIleGlyArgThrGlyArg 619

Qy        1900 TGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATTTAGCA 1959
               |||||||||||||||:::|||   |||||||||:::::::::   |||:::   :::|||
Db         620 CysGlyAsnThrGlyLysAlaThrSerPhePheAsnValGlnAspAspHisValIleAla 639

Qy        1960 CAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAA 2019
               :::||||||||||||:::|||||||||||||   |||:::||||||||||||||||||||||
Db         640 ArgProLeuValLysIleLeuThrAspAlaHisGlnGluValProAlaTrpLeuGluGlu 659

Qy        2020 ATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCA 2079
               ||||||                     |||   |||         |||   |||
Db         660 IleAla---------------------PheGlyGlyHisGlyAlaLeuAsnSerPheTyr 672

Qy        2080 TCAGTTGATACC---------AGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCT 2130
               :::   |||:::         :::   |||   :::   :::   |||      |||:::
Db         673 AlaAlaAspSerMetGlyGluGlnAlaGlyGlyAsnAlaValThrThrProSerPheAla 692

Qy        2131 TCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
                                         ::::::   |||||||||
Db         693 GlnGlu-----------------GluGluAlaSerTrpAsp 700
```

```
RESULT 5
A58768
ATP-dependent RNA helicase homolog - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 31-Dec-1990 #sequence_revision 17-Apr-1998 #text_change 19-Jan-2001
C;Accession: A58768; S01676; S10129; A31922
R;Liang, L.; Diehl-Jones, W.; Lasko, P.
unpublished results, 1995, cited by GenBank in release 103.0
A;Description: Localization of Vasa protein to the Drosophila pole plasm is independent
A;Reference number: A58768
A;Accession: A58768
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-661 <LAS1>
A;Cross-references: EMBL:X12945; NID:g433675; PIDN:CAA31405.1; PID:g1054723
R;Lasko, P.F.; Ashburner, M.
Nature 335, 611-617, 1988
A;Title: The product of the Drosophila gene vasa is very similar to eukaryotic initiatio
A;Reference number: S01676; MUID:89014721; PMID:3140040
A;Accession: S01676
A;Molecule type: DNA
A;Residues: 1-34,'R',36-448,'LRF',453-460,'R',462-589,'Q',591-661 <LAS2>
A;Cross-references: EMBL:X12945; NID:g433675
R;Ashburner, M.
submitted to the EMBL Data Library, October 1988
A;Reference number: S10129
A;Accession: S10129
A;Molecule type: DNA
A;Residues: 1-34,'R',36-644,'Q',645-661 <ASH>
A;Cross-references: EMBL:X12945; NID:g433675
R;Hay, B.; Jan, L.Y.; Jan, Y.N.
Cell 55, 577-587, 1988
A;Title: A protein component of Drosophila polar granules is encoded by vasa and has exte
A;Reference number: A31922; MUID:89028669; PMID:3052853
A;Accession: A31922
A;Molecule type: mRNA
A;Residues: 1-34,'R',36-154,168-264,'Y',266-321,'C',323-451,'F',453-581,'R',583-593,'H',5
A;Cross-references: GB:M23560; NID:g158795; PIDN:AAA29013.1; PID:g158796
A;Note: the authors translated the codon TGT for residue 322 as Val
C;Genetics:
A;Gene: vasa
A;Cross-references: FlyBase:FBgn0003970
A;Introns: 8/3; 138/1; 234/3; 482/3; 554/1; 644/2
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;289-296/Region: nucleotide-binding motif A (P-loop)
F;395-400/Region: nucleotide-binding motif B
F;399-402/Region: DEAD motif

Alignment Scores:
Pred. No.:              2.15e-83      Length:          661
Score:                  1274.00       Matches:         314
Percent Similarity:     56.73%        Conservative:    103
Best Local Similarity:  42.72%        Mismatches:      228
Query Match:            32.22%        Indels:          90
DB:                     2             Gaps:            25

US-09-714-865-15 (1-2172) x A58768 (1-661)

Qy          10 GAAGATTGGGAAGCAGAAATCAACCCTCATATGTCTTCCTATGTTCCCATATTTGAG--- 66
               :::||||||:::   |||                           ||||||   :::
Db           3 AspAspTrpAspAspGlu---------------------------ProIleValAspThr 13

Qy          67 AAGGATAGGTATTCTGGAGAAAATGGAGACAATTTTAACAGGACTCCAGCTTCATCATCA 126
               :::            |||:::      |||:::   :::         :::   |||
Db          14 ArgGlyAlaArgGlyGlyAspTrpSerAspAspGluAspThrAlaLysSerPheSerGly 33

Qy         127 GAAATGGAT---GATGGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCT 183
               |||   :::   ||||||                              |||
Db          34 GluAlaGluGlyAspGlyValGlyGly---------------SerGlyGlyGluGlyGly 48

Qy         184 GGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCCACA 243
               |||      ||||||||||||
Db          49 GlyTyrGlnGlyGlyAsnArgAsp--------------------------------Val 57

Qy         244 ATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTT 303
                  |||      |||   |||:::      |||      |||:::
Db          58 PheGlyArgIleGlyGlyGlyArgGlyGlyGlyAlaGlyGlyTyrArgGlyGlyAsnArg 77

Qy         304 GAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCA 363
               :::   |||         |||      ||||||         |||
Db          78 AspGlyGlyGlyPheHisGlyGlyArgArgGluGlyGluArgAspPhe------------ 93

Qy         364 ACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCT 423
                        ||||||               ||||||:::|||   |||
Db          94 ---------ArgGly------GlyGluGlyGlyPheArgGlyGlyGlnGlyGlySerArg 108

Qy         424 TCAGGGCCA---TACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
                  |||         |||||||||:::|||   |||||||||      |||||||||
Db         109 GlyGlyGlnGlyGlySerArgGlyGlyGlnGlyGlyPheArgGlyGlyGluGlyGlyPhe 128
```

```
Qy      481 ---GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGT 537
               |||          |||  |||  |||                 :::|||  |||
Db      129 ArgGlyArgLeuTyrGluAsnGluAspGlyAsp---------------GluArgArgGly 143

Qy      538 GGCCTT---------------TTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAAT 582
               |||                   |||  ||||||    |||
Db      144 ArgLeuAspArgGluGluArgGlyGlyGluArgArgGlyArgLeuAsp------------ 159

Qy      583 GGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGA------GGTGGTTAC--- 633
                     :::  |||  |||         ||||||    ||||||:::
Db      160 ---------ArgGluGluArgGlyGlyGluArgGlyGluArgGlyAspGlyGlyPheAla 176

Qy      634 ---AAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAA 690
               :::       ||||||:::   |||                |||:::    :::
Db      177 ArgArgArgArgAsnGluAspAspIle---------------AsnAsnAsnAsnAsnIle 191

Qy      691 GCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCT 750
            ||||||               |||   :::    :::      ||||||||||||
Db      192 AlaGlu---------------AspValGluArgLysArgGluPheTyrIleProProGlu 206

Qy      751 CCACCTGAGGATGAGGACTCCATCTTTGCACAT---TATCAGACAGGCATAAACTTCGAC 807
            |||    |||     ||||||:::            :::||||||:::|||
Db      207 ProSerAsnAspAlaIleGluIlePheSerSerGlyIleAlaSerGlyIleHisPheSer 226

Qy      808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
            ||||||:::   |||   |||:::|||:::|||   |||   |||      |||
Db      227 LysTyrAsnAsnIleProValLysValThrGlySerAspValProGlnProIleGlnHis 246

Qy      868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
            |||      |||:::|||         :::   :::|||:::   |||:::|||:::
Db      247 PheThrSerAlaAspLeuArgAspIleIleIleAspAsnValAsnLysSerGlyPheLys 266

Qy      928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
               ||||||:::||||||   |||||||||:::|||   :::||||||||||||||
Db      267 IleProThrProIleGlnLysCysSerIleProValIleSerSerGlyArgAspLeuMet 286

Qy      988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db      287 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerLys 306

Qy     1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
            ::::::  |||                  ||||||    |||:::   :::
Db      307 LeuLeuGluAsp------------ProHisGluLeuGluLeuGlyArgProGlnValVal 322

Qy     1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
            ||||||:::||||||||||||||      ||||||:::   ||||||||||||||:::
Db      323 IleValSerProThrArgGluLeuAlaIleGlnIlePheAsnGluAlaArgLysPheAla 342

Qy     1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
            |||   :::   ::::::       ::::::|||||||||||       |||
Db      343 PheGluSerTyrLeuLysIleGlyIleValTyrGlyGlyThrSerPheArgHisGlnAsn 362

Qy     1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
                |||   :::||||||:::::::::   ||||||||||||||||:::|||
Db      363 GluCysIleThrArgGlyCysHisValValIleAlaThrProGlyArgLeuLeuAspPhe 382

Qy     1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
            :::   :::     |||     :::     :::::::::||||||||||||||||
Db      383 ValAspArgThrPheIleThrPheGluAspThrArgPheValValLeuAspGluAlaAsp 402

Qy     1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
            ||||||||||||||||||        :::|||:::::::::::::::        |||
Db      403 ArgMetLeuAspMetGlyPheSerGluAspMetArgArgIleMetThrHisValThrMet 422

Qy     1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
                  :::   :::   |||||||||||||||||||||||||||||||||||||||
Db      423 ------ArgProGluHisGlnThrLeuMetPheSerAlaThrPheProGluGluIleGln 440

Qy     1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGT 1527
            |||:::|||   ||||||||||||   ||||||:::   ||||||:::|||   ||||||
Db      441 ArgMetAlaGlyGluPheLeuLys---AsnTyrValSerValAlaIleGlyIleValGly 459

Qy     1528 GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAA 1587
            |||||||||   ||||||:::||||||:::   :::|||   ::::::::::||||||
Db      460 GlyAlaCysSerAspValLysGlnThrIleTyrGluValAsnLysTyrAlaLysArgSer 479

Qy     1588 AAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACT 1647
            ||||||:::|||||||||            |||         |||:::||||||||||||||
Db      480 LysLeuIleGluIleLeuSerGluGlnAlaAsp---GlyThrIleValPheValGluThr 498

Qy     1648 AAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGT 1707
            |||:::    |||||||||   |||:::|||||||   :::::::::     ||||||||
Db      499 LysArgGlyAlaAspPheLeuAlaSerPheLeuSerGluLysGluPheProThrThrSer 518

Qy     1708 ATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGA 1767
            |||||||||||||||   |||   :::|||||||||||||||   ||||||:::   |||
Db      519 IleHisGlyAspArgLeuGlnSerGlnArgGluGlnAlaLeuArgAspPheLysAsnGly 538

Qy     1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
                   ||||||:::||||||||||||||:::||||||||||||||||:::|||:::
Db      539 SerMetLysValLeuIleAlaThrSerValAlaSerArgGlyLeuAspIleLysAsnIle 558

Qy     1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
            :::||||||||||||:::|||:::||||||   ||||||:::||||||||||||||||||
Db      559 LysHisValIleAsnTyrAspMetProSerLysIleAspAspTyrValHisArgIleGly 578

Qy     1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
            |||||||||      ||||||   ||||||||||||   |||||||||||   |||   |||
Db      579 ArgThrGlyCysValGlyAsnAsnGlyArgAlaThrSerPhePheAspProGluLysAsp 598

Qy     1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCA 2007
                  :::|||      |||||||||:::|||      :::   |||   ||||||
Db      599 ArgAlaIleAlaAlaAspLeuValLysIleLeuGluGlySerGlyGlnThrValProAsp 618

Qy     2008 TGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
            :::|||                |||         |||      |||
Db      619 PheLeuArg---------------ThrCysGlyAlaGlyGlyAspGlyGlyTyrSerAsn 633

Qy     2068 AACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
                  |||   ||||||   |||   |||:::         |||      |||
Db      634 GlnAsnPheGlyGlyValAspValArg---GlyArgGly------AsnTyrValGly--- 649

Qy     2128 TCTTCTTCACGAGCTCCCAATCCAGTAGATGATGAGTCATGGGAT 2172
                        |||   |||      :::::|||   ||||||
Db      650 ---------AspAlaThrAsnValGluGluGluGlnTrpAsp 661
```

RESULT 6
I84741
RNA helicase - mouse
N;Alternate names: RNA helicase ERH
C;Species: Mus musculus (house mouse)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 19-Jan-2001
C;Accession: I84741; S56112
R;Gee, S.L.; Conboy, J.G.
Gene 140, 171-177, 1994
A;Title: Mouse erythroid cells express multiple putative RNA helicase genes exhibiting h
A;Reference number: I49731; MUID:94192995; PMID:8144024
A;Accession: I84741
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-662 <RES>
A;Cross-references: GB:L25126; NID:g407995; PIDN:AAA53630.1; PID:g407996
A;Genetics: RES1
R;Sowden, J.; Putt, W.; Morrison, K.; Beddington, R.; Edwards, Y.
Biochem. J. 308, 839-846, 1995
A;Title: The embryonic RNA helicase gene (ERH): a new member of the DEAD box family of RN
A;Reference number: S56112; MUID:97104282; PMID:8948440
A;Accession: S56112
A;Molecule type: mRNA
A;Residues: 1-229,'R',231-662 <SOW>
A;Cross-references: GB:Z38117; NID:g1835121; PID:g1835122
A;Experimental source: strain c57bl/6; notochord
A;Genetics: SOW1
A;Note: the sequence is revised in GenBank entry MMDBRNAHL, release 117, (PID:1835122)
A;Note: the revised sequence is now identical to PIR accession I84741
C;Genetics: <RES1>

```
A;Gene: Eif4a-rs1; MGI:Ddx19
A;Cross-references: MGI:99526
C;Genetics: <SOW1>
A;Gene: MGI:Ddx3
A;Cross-references: MGI:103064
A;Map position: 1
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;224-231/Region: nucleotide-binding motif A (P-loop)
F;343-348/Region: nucleotide-binding motif B
F;347-350/Region: DEAD motif

Alignment Scores:
Pred. No.:              3.63e-80     Length:        662
Score:                  1229.00      Matches:       297
Percent Similarity:     57.59%       Conservative:  90
Best Local Similarity:  44.20%       Mismatches:    183
Query Match:            31.08%       Indels:        102
DB:                     1            Gaps:          22

US-09-714-865-15 (1-2172) x I84741 (1-662)

Qy    214 GAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGT--- 270
Db     20 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 37

Qy    271 ---------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
Db     38 TyrIleProProHisLeuArgAsnArgGlu---AlaThrLysGlyPheTyrAspLysAsp 56

Qy    316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGA 375
Db     57 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSer--------------- 71

Qy    376 GGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATAC 435
Db     72 ---PheGlySerArgGlyAspSerArg------GlyLysSerSerPhePheGlyAspArg 88

Qy    436 AGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGCCGTGGAGGATTT---GGTCTA 486
Db     89 GlySerGlySerArgGlyArgPheAspAspArgGly---ArgGlyAspTyrAspGlyIle 107

Qy    487 GGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTT 546
Db    108 Gly--------------------------------------------------------- 108

Qy    547 GGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGT 606
Db    109 ------------------------------GlyArgGlyAsp------------ArgSer 114

Qy    607 GGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGC 663
Db    115 GlyPheGlyLysPheGluArgGlyGly------------AsnSerArgTrpCysAspLys 130

Qy    664 TCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGA 723
Db    131 SerAspGluAspAspTrpSerLys------------------------------------ 138

Qy    724 CCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT---GAGGACTCCATCTTTGCA 780
Db    139 ------------------ProLeuProProSerGluArgLeuGluGlnGluLeuPheSer 152

Qy    781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
Db    153 GlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIleProValGluAlaThrGly 172

Qy    841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
Db    173 AsnAsnCysProProHisIleGluSerPheSerAspValGluMetGlyGluIleIleMet 192

Qy    901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
Db    193 GlyAsnIleGluLeuThrArgTyrThrArgProThrProValGlnLysHisAlaIlePro 212

Qy    961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
Db    213 IleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 232

Qy   1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA------------ATAACT 1068
Db    233 AlaPheLeuLeuProIleLeuSerGlnIleTyrAlaAspGlyProGlyGluAlaLeuArg 252

Qy   1069 GCCAGTCGT---------------TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTA 1113
Db    253 AlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyrProIleSerLeuValLeu 272

Qy   1114 GCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGG 1173
Db    273 AlaProThrArgGluLeuAlaValGlnIleTyrGluGluAlaArgLysPheSerTyrArg 292

Qy   1174 ACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAA 1233
Db    293 SerArgValArgProCysValValTyrGlyGlyAlaGluIleGlyGlnGlnIleArgAsp 312

Qy   1234 ATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGC 1293
Db    313 LeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArgLeuValAspMetMetGlu 332

Qy   1294 AAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATG 1353
Db    333 ArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeuAspGluAlaAspArgMet 352

Qy   1354 TTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCA 1413
Db    353 LeuAspMetGlyPheGluProGlnIleArgArgIleValGluGlnAspThrMetProPro 372

Qy   1414 AAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTG 1473
Db    373 LysGlyValArgHisThrMetMetPheSerAlaThrPheProLysGluIleGlnMetLeu 392

Qy   1474 GCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCA 1533
Db    393 AlaArgAspPheLeu---AspGluTyrIlePheLeuAlaValGlyArgValGlySerThr 411

Qy   1534 TGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTC 1593
Db    412 SerGluAsnIleThrGlnLysValValTrpValGluGluIleAspLysArgSerPheLeu 431

Qy   1594 GTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATGGTCTTTGTTGAAACTAAG 1650
Db    432 LeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeuValPheValGluThrLys 451

Qy   1651 AAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATC 1710
Db    452 LysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGlyTyrAlaCysThrSerIle 471

Qy   1711 CATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAG 1770
Db    472 HisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHisGlnPheArgSerGlyLys 491

Qy   1771 TGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAA 1830
Db    492 SerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeuAspIleSerAsnValLys 511

Qy   1831 CATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGT 1890
Db    512 HisValIleAsnPheAspLeuProSerAspIleGluGluTyrValHisArgIleGlyArg 531

Qy   1891 ACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAAC 1950
Db    532 ThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePheAsn---GluArgAsnIle 550

Qy   1951 CATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGG 2010
Db    551 AsnIleThrLysAspLeuLeuAspLeuLeuValGluAlaLysGlnGluValProSerTrp 570
```

```
Qy       2011 TTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAAC 2070
              |||||   :::|||||       :::          :::  ||||||:::||||||
Db        571 LeuGluAsnMetAlaPheGluHisHis---------TyrLysGlySerSerArgGlyArg 587

Qy       2071 GTG-----------------------TTTGCATCAGTTGATACCAGAAAGGGCAAGAGC 2106
                                         |||   :::   |||   |||:::      |||
Db        588 SerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAspTyrArgGln---SerSer 606

Qy       2107 ACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                    :::::::::   ||||||||||||||||
Db        607 GlyAlaSerSerSerSerPheSerSerSerArgAla 618

RESULT 7
S13654
ATP-dependent RNA helicase - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: S13654; S29676
R;Gururajan, R.; Perry-O'Keefe, H.; Melton, D.A.; Weeks, D.L.
Nature 349, 717-719, 1991
A;Title: The Xenopus localized messenger RNA An3 may encode an ATP-dependent RNA helicas
A;Reference number: S13654; MUID:91141586; PMID:1996140
A;Accession: S13654
A;Molecule type: mRNA
A;Residues: 1-697 <GUR>
A;Cross-references: EMBL:X57328; NID:g65059; PIDN:CAA40605.1; PID:g65060
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;260-272/Region: ATP binding #status predicted
F;265-272/Region: nucleotide-binding motif A (P-loop)
F;384-389/Region: nucleotide-binding motif B
F;386-398/Region: ATP binding #status predicted
F;388-391/Region: DEAD motif

Alignment Scores:
Pred. No.:              7.74e-80      Length:          697
Score:                  1224.50       Matches:         290
Percent Similarity:     56.85%        Conservative:    96
Best Local Similarity:  42.71%        Mismatches:      198
Query Match:            30.97%        Indels:          95
DB:                     1             Gaps:            18

US-09-714-865-15 (1-2172) x S13654 (1-697)

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
                :::   ||||||       |||   |||:::|||
Db         47 AlaSerArgAsnAspSerAsnTrpAspSerGlyArg------------------------ 58

Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
                                                |||   |||:::   |||
Db         59 ---------------------------------GlyGlyAsnGlyTyrIleAsnGlyMet 67

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAAT 360
                 :::|||   |||                |||   :::      |||
Db         68 GlnAspAspArgAspGly------------ArgMetAsnGlyTyrAspArgGlyGlyTyr 83

Qy        361 CCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAA 420
                 :::|||      |||   |||   ||||||   |||   |||      ||||||
Db         84 GlySerArgGlyThrGlyArgSerAspArgGlyPheTyr---AspArgGluAsnSerGly 102

Qy        421 GCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTT 480
                 :::          :::               ||||||   |||   ||||||
Db        103 TrpAsnSerGlyArgAspLysAspAlaTyrSerSerPhe---GlySerArgGlyAspArg 121

Qy        481 GGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGC 540
              |||   ||||||
Db        122 GlyLysGlySer------------------------------------------------ 125

Qy        541 CTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGC 600
              ||||||   |||               |||:::|||:::         |||
Db        126 LeuPheAsnGluArg---------------GlySerGlySerArgArgThrAspAspArg 140

Qy        601 AGAAGT------------------GGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTA 642
              |||                     |||:::   |||:::::   |||:::
Db        141 ArgGlnAspGlyPheAspGlyMetGlyAsnArgSerAspLysSerGlyPheGlyArgPhe 160

Qy        643 AATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGA 702
              :::                        |||   :::   |||
Db        161 AspArg------------------GlyAsnSerArgTrp--------------------- 167

Qy        703 GAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT 762
                    ||||||   :::      :::         :::      |||      |||   |||
Db        168 ------SerAspAspArgAsnAspGluAspAspTrpSerLysProLeuAlaProAsnAsp 185

Qy        763 ------GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGAC 816
                    |||      :::|||:::         |||||||||||||||:::|||||||||
Db        186 ArgValGluGlnGluLeuPheSerGlySerAsnThrGlyIleAsnPheGluLysTyrAsp 205

Qy        817 ACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAA 876
                 |||   ||||||   :::|||   :::   ||||||   |||   :::|||   :::
Db        206 AspIleProValGluAlaThrGlySerAsnCysProProHisIleGluSerPheHisAsp 225

Qy        877 GCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACT 936
                 :::   :::   :::         ||||||            ||||||:::   |||
Db        226 ValThrMetGlyGluIleIleMetGlyAsnIleGlnLeuThrArgTyrThrArgProThr 245

Qy        937 CCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCT 996
              |||||||||||||||::::::|||||||||||||:::         ||||||||||||||||||
Db        246 ProValGlnLysHisAlaIleProIleIleIleGluLysArgAspLeuMetAlaCysAla 265

Qy        997 CAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTGGCTCATATGATGCAT 1056
              ||||||||||||||||||||||||||||||||||||||||||:::   :::
Db        266 GlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrAla 285

Qy       1057 GATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA-------------------- 1095
              ||||||                  |||   |||||||||
Db        286 AspGly---ProGlyAspAlaMetLysHisLeuGlnGluAsnGlyArgTyrGlyArgArg 304

Qy       1096 ---------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
                       |||      ::::::::||||||||||||||||||      |||||||||
Db        305 LysGlnPheProLeuSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyr 324

Qy       1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
                 |||||||||||||||::::::   :::   ||||||      |||:::|||||||||
Db        325 GluGluAlaArgLysPheAlaTyrArgSerArgValArgProCysValValTyrGlyGly 344

Qy       1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
                 :::|||      ||||||   :::   :::||||||:::::::|||   |||
Db        345 AlaAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThr 364

Qy       1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
              ||||||||||||:::|||::::::   :::   ||||||||||||         ||||||
Db        365 ProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyr 384

Qy       1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
              ||||||||||||||||||||||||||||||||||||||||||   |||::::::::::::
Db        385 LeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArg 404

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              ::::::            ||||||   |||      |||||||||:::||||||||||||
Db        405 IleValGluGlnAspThrMetProProLysGlyValArgGlnThrMetMetPheSerAla 424

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              |||||||||:::|||||||||   ||||||      :::||||||      |||:::|||
Db        425 ThrPheProLysGluIleGlnIleLeuAlaArgAspPheLeu---AspGluTyrIlePhe 443

Qy       1507 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1566
              :::|||||||||:::||||||               ::::::   |||   |||:::   |||
Db        444 LeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpVal 463

Qy       1567 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAA 1623
                 :::      ||||||      |||:::::::::|||         |||   |||
Db        464 GluGluMetAspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSer 483

Qy       1624 AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGT 1683
                 |||:::|||||||||||||||||||||   ||||||            ||||||
```

```
Db        484 LeuThrLeuValPheValGluThrLysLysGlyAlaAspAlaLeuGluAspPheLeuTyr 503
Qy       1684 CAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAA 1743
Db        504 HisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGlu 523
Qy       1744 GCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCC 1803
Db        524 AlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAla 543
Qy       1804 AGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATT 1863
Db        544 ArgGlyLeuAspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIle 563
Qy       1864 GATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATT 1923
Db        564 GluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThr 583
Qy       1924 TCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACA 1983
Db        584 SerPhePheAsn---GluLysAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuVal 602
Qy       1984 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCT 2043
Db        603 GluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaTyrGluGlnHis------ 620
Qy       2044 GGCTTCAGTGGTAGTACAAGAGGAAACGTG--------------------TTTGCATCA 2082
Db        621 ---HisLysSerSerSerArgGlyArgSerLysSerArgPheSerGlyGlyPheGlyAla 639
Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
Db        640 LysAspTyrArgGlnSerSerGlyAla------GlySerSerPheGlySerSerArg 656

RESULT 8
A32378
probable ATP-dependent RNA helicase - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Sep-1989 #sequence_revision 08-Sep-1989 #text_change 19-Jan-2001
C;Accession: A32378
R;Leroy, P.; Alzari, P.; Sassoon, D.; Wolgemuth, D.; Fellous, M.
Cell 57, 549-559, 1989
A;Title: The protein encoded by a murine male germ cell-specific transcript is a putativ
A;Reference number: A32378; MUID:89249320; PMID:2720782
A;Accession: A32378
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-660 <LER>
A;Cross-references: GB:J04847; NID:g200388; PIDN:AAA39942.1; PID:g200389
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; DNA binding; nucleotide binding; P-loop
F;223-230/Region: nucleotide-binding motif A (P-loop)
F;342-347/Region: nucleotide-binding motif B
F;346-349/Region: DEAD motif

Alignment Scores:
Pred. No.:              7.17e-78       Length:        660
Score:                  1197.00        Matches:       286
Percent Similarity:     57.18%         Conservative:  104
Best Local Similarity:  41.94%         Mismatches:    186
Query Match:            30.27%         Indels:        106
DB:                     2              Gaps:          19

US-09-714-865-15 (1-2172) x A32378 (1-660)

Qy        181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGATAATACATCC 240
Db         17 AlaGlyLeuAspLeuThrSerArgAspSer---------GlnSerGlyGlySerThrAla 33
Qy        241 ACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTTGGAAACAGAGGTTTTTCAAACAGCAGG 300
Db         34 SerLysGlyArgTyr---IleProProHisLeuArgAsnArgGluAlaAlaLysAla--- 51
Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAG---------------TCTAGTAAT 345
Db         52 PheTyrAspLysAspGlySerArgTrpSerLysAspLysAspAlaTyrSerSerPheGly 71
Qy        346 GACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGAT 405
Db         72 SerArgSerAspThrArgAlaLysSerSerPhePheSerAspArgGlyGly--------- 88
Qy        406 GGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGA--- 462
Db         89 ---------SerGlySerArgGlyArgPheAspGluArgGlyArgSerAspTyrGluSer 105
Qy        463 ---GGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAA 519
Db        106 ValGlySerArgGlyGly------------------------------------------ 111
Qy        520 TGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGT 579
Db        111 ------------------------------------------------------------ 111
Qy        580 AATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTACAAA 636
Db        112 --------------------ArgSerGlyPheGlyLysPheGluArgGly--------- 121
Qy        637 GGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAA 696
Db        122 -----------------------------GlyAsnSerArgTrpCysAspLysAlaAsp 131
Qy        697 GGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCT 756
Db        132 GluAspAspTrpSerLys---------------------------ProLeuProProSer 142
Qy        757 GAGGAT---GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATAC 813
Db        143 GluArgLeuGluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyr 162
Qy        814 GACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAA 873
Db        163 AspAspIleProValGluAlaThrGlyAsnAsnCysProProHisIleGluSerPheSer 182
Qy        874 GAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTT 933
Db        183 AspValGluMetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgPro 202
Qy        934 ACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGT 993
Db        203 ThrProValGlnLysHisAlaIleProIleIleLysGluLysArgAspLeuMetAlaCys 222
Qy        994 GCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATG 1053
Db        223 AlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyr 242
Qy       1054 CATGATGGA------------ATAACTGCCAGTCGT---------------TTTAAAGAG 1086
Db        243 ThrAspGlyProGlyGluAlaLeuArgAlaMetLysGluAsnGlyLysTyrGlyArgArg 262
Qy       1087 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
Db        263 LysGlnTyrProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyr 282
Qy       1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
Db        283 GluGluAlaArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGly 302
Qy       1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
Db        303 AlaAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThr 322
Qy       1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
Db        323 ProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyr 342
Qy       1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
```

```
                 |||||||||||||||||||||||||||||||||||||||||    |||:::::::::::
Db        343 LeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArg 362

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              ::::::              ||||||   |||        |||    |||:::|||||||||||
Db        363 IleValGluGlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAla 382

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              ||||||||:::||||||||     ||||||    :::||||||       |||:::|||
Db        383 ThrPheProLysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePhe 401

Qy       1507 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1566
              :::||||||||:::||||||             ::::::    |||   |||:::    |||
Db        402 LeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpVal 421

Qy       1567 GGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAA 1623
                 :::          ||||||     |||:::::::::|||          |||   |||
Db        422 GluGluAlaAspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSer 441

Qy       1624 AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTACTGCAACTTTTCTTTGT 1683
                    :::||||||||||||||||||||      ||||||               ||||||
Db        442 LeuIleLeuValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyr 461

Qy       1684 CAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAA 1743
                 |||       :::      ||||||||||||||||||||     ||||||:::||||||:::
Db        462 HisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGlu 481

Qy       1744 GCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCC 1803
              ||||||      ||||||    ||||||    |||:::|||||||||||||:::|||||||||
Db        482 AlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAla 501

Qy       1804 AGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATT 1863
              ||||||||||||||        ||||||:::||||||||||||||||||||||||||      |||
Db        502 ArgGlyLeuAspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIle 521

Qy       1864 GATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATT 1923
              :::|||||||||||||||||||||||||||||||||     ||||||    |||    |||
Db        522 GluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThr 541

Qy       1924 TCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACA 1983
              |||||||||:::      |||   :::   ::::::    :::    |||:::    :::|||
Db        542 SerPhePheAsn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuVal 560

Qy       1984 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCT 2043
              :::|||:::|||:::||||||:::|||||||||   :::||||||         :::
Db        561 GluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaPheGluHisHis------ 578

Qy       2044 GGCTTCAGTGGTAGTACAAGAGGAAACGTG--------------------TTTGCATCA 2082
                 :::      |||    :::||||||                   |||     :::
Db        579 ---TyrLysGlyGlySerArgGlyArgSerLysSerArgPheSerGlyGlyPheGlyAla 597

Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                 |||   |||:::       |||      :::::::::    |||||||||    ||||||
Db        598 ArgAspTyrArgGln---SerSerGlyAlaSerSerSerSerPheSerSerGlyArgAla 616

Qy       2143 CCCAAT 2148
                 |||
Db        617 SerAsn 618
```

RESULT 9
T48796
probable ATP-dependent RNA helicase DED1 (imported) - Neurospora crassa
N;Alternate names: protein 15E6.40
C;Species: Neurospora crassa
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 17-Nov-2000
C;Accession: T48796
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24541
A;Accession: T48796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-688 <SCH>
A;Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15E6.40
A;Experimental source: cosmid contig 15E6; strain 74
C;Genetics:
A;Gene: NCSP:15E6.40
A;Map position: 2
A;Introns: 63/2
C;Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:

Pred. No.:	2.33e-76	Length:	688
Score:	1176.00	Matches:	265
Percent Similarity:	60.74%	Conservative:	80
Best Local Similarity:	46.65%	Mismatches:	177
Query Match:	29.74%	Indels:	46
DB:	2	Gaps:	11

US-09-714-865-15 (1-2172) x T48796 (1-688)

```
Qy        409 AATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGC 468
              ||||||    :::|||    ||||||            ||||||    |||::::::
Db         65 AsnAsnTyrAspAlaArgGlyPro---------GlyGlyGlyGlyAsnTrpProAlaPro 81

Qy        469 CGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG 528
                 |||      |||   |||
Db         82 GlyGlyProProGlyProGlyPheGluGly----------------------------- 91

Qy        529 CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGAT 588
                 :::   |||      :::|||        ||||||         |||     |||   :::
Db         92 GlnGlnGlyAlaGlyTrpGlyGlyProArgPro------GlnGlyGlyPheAsnProAsn 109

Qy        589 ACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAA 648
                 :::        :::|||:::|||:::         |||          |||
Db        110 AlaTyrArgGlyAsnAlaGlyAlaGlyAlaGlyAlaGlyAlaGlyGlyGlyGlyGlySer 129

Qy        649 GAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGT 708
                       |||||||||       |||:::             :::|||
Db        130 PheSerAsnArgGlySerGlyAspGlyGlnTrpArg---------AspGly--------- 143

Qy        709 AGTGATACTCAAGGACCAAAAGTGACCTACATACCC---CCTCCTCCACCTGAGGATGAG 765
                                      :::||||||   |||      |||      |||
Db        144 -----------------------LysHisIleProGlyProAlaAsnProArgValGlu 155

Qy        766 GACTCCATCTTT------------------GCACATTATCAGACAGGCATAAACTTCGAC 807
                 :::|||                   :::         |||||||||||||||:::
Db        156 ArgGluLeuPheGlyThrProAspAspProSerLysGlnHisThrGlyIleAsnPheGlu 175

Qy        808 AAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACT 867
              |||||||||   |||   ||||||   ||||||   :::      |||      :::||||||
Db        176 LysTyrAspAspIleProValGluAlaSerGlyAspAsnValProGluProValLeuThr 195

Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
               |||         |||         |||   :::||||||       |||    |||
Db        196 PheSerAsnProProLeuAspAsnHisLeuIleSerAsnIleGlnLeuAlaArgTyrAsn 215

Qy        928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
                 ||||||||||||||||||||||||||||:::::::   |||||||||||||||
Db        216 ValProThrProValGlnLysTyrSerIleProIleValMetGlyGlyArgAspLeuMet 235

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              |||||||||||||||||||||||||||||         ||||||    |||||||||:::
Db        236 AlaCysAlaGlnThrGlySerGlyLysThrGlyGlyPheLeuPheProIleLeuSerGln 255

Qy       1048 ATGATGCATGATGGA---------ATAACTGCCAGTCGTTTT--------------AAA 1083
                 |||   |||       |||   ||||||                             :::
Db        256 SerPheHisThrGlyProSerProIleProAlaSerAlaAlaGlyAlaTyrGlyArgGln 275

Qy       1084 GAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATT 1143
                    |||         :::|||:::|||||||||||||||||||||||:::||||||
Db        276 ArgLysAlaTyrProThrAlaLeuIleLeuAlaProThrArgGluLeuValSerGlnIle 295

Qy       1144 TATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGG 1203
              |||   ||||||||||||||:::::::   :::   ||||||      |||:::||||||
Db        296 TyrAspGluAlaArgLysPheAlaTyrArgSerTrpValArgProCysValValTyrGly 315
```

```
Qy       1204 GGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCT 1263
              |||      :::|||      :::||||||||    :::||||||::::::|||    |||
Db        316 GlyAlaAspIleGlySerGlnLeuArgGlnIleGluArgGlyCysAspLeuLeuValAla 335

Qy       1264 ACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAA 1323
              ||||||||||||||:::|||:::|||   :::   :::|||   |||      ||||||
Db        336 ThrProGlyArgLeuValAspLeuIleGluArgGlyArgIleSerLeuCysAsnIleLys 355

Qy       1324 TACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAG 1383
              ||||||||||||||||||||||||||||||||||||||||||||    |||:::::::::
Db        356 TyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArg 375

Qy       1384 AAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGT 1443
              :::::::::           ||||||           ||||||||||||||||||
Db        376 ArgIleValGluGlyGluAspMetProLysValAsnAspArgGlnThrLeuMetPheSer 395

Qy       1444 GCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTG 1503
              ||||||||||||    :::||||||    ||||||    :::||||||||||   :::|||:::
Db        396 AlaThrPheProArgAspIleGlnIleLeuAlaArgAspPheLeuLys---AspTyrIle 414

Qy       1504 TTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAA 1563
              |||::::::||||||:::||||||              ::::::   |||   |||
Db        415 PheLeuSerValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValGluTyr 434

Qy       1564 GTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAA 1623
              |||                  ||||||      |||::::::||||||
Db        435 ValGluAspIleAspLysArgSerValLeuLeuAspIleLeuHisThrHisAlaGlyGly 454

Qy       1624 AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGT 1683
                 |||::::::|||||||||||||||:::     ||||||       :::    ||||||
Db        455 LeuThrLeuIlePheValGluThrLysArgMetAlaAspSerLeuSerAspPheLeuIle 474

Qy       1684 CAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAA 1743
                 :::              ||||||||||||||||||   ||||||||||||||:::
Db        475 AsnGlnAsnPheProAlaThrSerIleHisGlyAspArgThrGlnArgGluArgGluArg 494

Qy       1744 GCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCC 1803
              ||||||      ||||||   |||:::||||||:::||||||||||||:::||||||||
Db        495 AlaLeuGluMetPheArgAsnGlyArgCysProIleLeuValAlaThrAlaValAlaAla 514

Qy       1804 AGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATT 1863
              ||||||||||||||   ||||||   ||||||||||||:::|||||||||:::    |||
Db        515 ArgGlyLeuAspIleProAsnValThrHisValIleAsnTyrAspLeuProThrAspIle 534

Qy       1864 GATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATT 1923
              |||:::||||||||||||||||||||||||||||||    ||||||||||||    |||
Db        535 AspAspTyrValHisArgIleGlyArgThrGlyArgAlaGlyAsnThrGlyIleAlaThr 554

Qy       1924 TCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACA 1983
              :::||||||:::            :::    :::   :::    |||::::::::::|||
Db        555 AlaPhePheAsn---ArgGlyAsnArgGlyValValArgGluLeuLeuGluLeuLeuLys 573

Qy       1984 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCT 2043
              :::|||    |||:::|||||||||:::||||||    ||||||        :::
Db        574 GluAlaAsnGlnGluValProAlaPheLeuGluThrIleAlaArgGluSerSerPheGly 593

Qy       2044 GGCTTCAGTGGTAGTACAAGAGGA 2067
              |||      |||         |||
Db        594 GlyGlyArgGlyGlyArgGlyGly 601
```

RESULT 10
S13653
ATP-dependent RNA helicase DED1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O4836; protein YOR204w
C;Species: Saccharomyces cerevisiae
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 19-Jan-2001
C;Accession: S13653; S07683; S67096
R;Jamieson, D.J.; Rahe, B.; Pringle, J.; Beggs, J.D.
Nature 349, 715-717, 1991
A;Title: A suppressor of a yeast splicing mutation (prp8-1) encodes a putative ATP-depen
A;Reference number: S13653; MUID:91141585; PMID:1996139
A;Accession: S13653
A;Molecule type: DNA
A;Residues: 1-604 <JAM>
A;Cross-references: EMBL:X57278; NID:g3646; PIDN:CAA40546.1; PID:g3647
A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 37-Met a
R;Struhl, K.
Nucleic Acids Res. 13, 8587-8601, 1985
A;Title: Nucleotide sequence and transcriptional mapping of the yeast pet56-his3-ded1 ger
A;Reference number: S07681; MUID:86093663; PMID:3001645
A;Accession: S07683
A;Molecule type: DNA
A;Residues: 1-36,'M',38-113 <STR>
A;Cross-references: EMBL:X03245; NID:g3778; PIDN:CAA27004.1; PID:g3781
R;Hughes, B.; Pohl, T.M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66685
A;Accession: S67096
A;Molecule type: DNA
A;Residues: 1-604 <HUG>
A;Cross-references: EMBL:Z75110; NID:g1420476; PIDN:CAA99419.1; PID:g1420479; MIPS:YOR20
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:DED1; SPP81
A;Cross-references: SGD:S0005730; MIPS:YOR204w
A;Map position: 15R
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;186-193/Region: nucleotide-binding motif A (P-loop)
F;302-307/Region: nucleotide-binding motif B
F;306-309/Region: DEAD motif

Alignment Scores:

Pred. No.:	8.02e-73	Length:	604
Score:	1126.50	Matches:	261
Percent Similarity:	56.66%	Conservative:	92
Best Local Similarity:	41.89%	Mismatches:	187
Query Match:	28.49%	Indels:	83
DB:	2	Gaps:	16

US-09-714-865-15 (1-2172) x S13653 (1-604)

```
Qy        337 TCTAGTAATGACTGCGAAGATAAT---------CCAACACGGAACAGAGGGTTTTCCAAG 387
              |||  ||||||      :::|||           |||      ||||||      :::
Db         12 SerIleAsnAspAsnAsnGluAsnGlyTyrValProProHisLeuArgGlyLysProArg 31

Qy        388 AGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGT--- 444
                                  |||   :::      :::|||   |||      |||
Db         32 SerAlaArgAsnAsnSerSerAsnTyrAsnAsnAsnAsnGlyGlyTyrAsnGlyGlyArg 51

Qy        445 GGAAGAGGTAGTTTC------CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAAT 498
              |||   |||||||||                  |||||||||:::|||
Db         52 GlyGlyGlySerPhePheSerAsnAsnArgArgGlyGlyTyrGly--------------- 66

Qy        499 AATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGA 558
                                                |||||||   ||||||
Db         67 ---------------------------------AsnGlyGlyPhePheGly--------- 72

Qy        559 CCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT 618
                          |||      |||      |||      |||||||||
Db         73 -----------GlyAsnAsnGlyGlySerArgSerAsnGlyArgSerGly--------- 85

Qy        619 GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT 678

Db         86 ---------------------------------------------------GlyArg 87

Qy        679 TGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTAC 738
              |||            :::|||                                :::
Db         88 Trp---------IleAspGly-------------------------------LysHis 93

Qy        739 ATACCCCCTCCTCCACCTGAGGAT---GAGGACTCCATCTTTGCA--------------- 780
              :::|||   |||      |||      |||   :::||||||
Db         94 ValProAlaProArgAsnGluLysAlaGluIleAlaIlePheGlyValProGluAspPro 113

Qy        781 CATTATCAG---ACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCT 837
```

```
                :::::|||    :::|||||||||||||    ||||||    |||    |||:::    |||
Db        114 AsnPheGlnSerSerGlyIleAsnPheAspAsnTyrAspAspIleProValAspAlaSer 133

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
                  |||   |||   |||         |||         |||                |||
Db        134 GlyLysAspValProGluProIleThrGluPheThrSerProProLeuAspGlyLeuLeu 153

Qy        898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
                     ||||||       |||    :::||||||   ||||||||||||||||||||:::
Db        154 LeuGluAsnIleLysLeuAlaArgPheThrLysProThrProValGlnLysTyrSerVal 173

Qy        958 CCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACT 1017
              ||||||:::       |||||||||||||||||||||||||||||||||||||||||||||
Db        174 ProIleValAlaAsnGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThr 193

Qy       1018 GCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACT--------- 1068
                    ||||||   |||:::|||:::                |||    :::
Db        194 GlyGlyPheLeuPheProValLeuSerGluSerPheLysThrGlyProSerProGlnPro 213

Qy       1069 ---------GCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCA 1119
                          |||   ::::::                |||      :::|||:::||||||
Db        214 GluSerGlnGlySerPheTyrGlnArgLysAlaTyrProThrAlaValIleMetAlaPro 233

Qy       1120 ACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGT 1179
              ||||||||||||      ||||||:::   ||||||:::||||||::::::      :::
Db        234 ThrArgGluLeuAlaThrGlnIlePheAspGluAlaLysLysPheThrTyrArgSerTrp 253

Qy       1180 GTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTA 1239
              |||:::|||   |||:::|||||||||:::    :::|||:::    :::|||:::|||
Db        254 ValLysAlaCysValValTyrGlyGlySerProIleGlyAsnGlnLeuArgGluIleGlu 273

Qy       1240 CAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAA 1299
              :::||||||::::::|||    ||||||||||||||||||    |||::::::    :::
Db        274 ArgGlyCysAspLeuLeuValAlaThrProGlyArgLeuAsnAspLeuLeuGluArgGly 293

Qy       1300 AAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGAT 1359
              ||||||   |||         :::|||||||||||||||||||||||||||||||||||||
Db        294 LysIleSerLeuAlaAsnValLysTyrLeuValLeuAspGluAlaAspArgMetLeuAsp 313

Qy       1360 ATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAA 1419
              |||||||||   |||:::::::::   ::::::            |||
Db        314 MetGlyPheGluProGlnIleArgHisIleValGluAspCysAspMetThrProValGly 333

Qy       1420 CAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCA 1479
              :::|||||||||||||||||||||||||||||||||      :::||||||   ||||||
Db        334 GluArgGlnThrLeuMetPheSerAlaThrPheProAlaAspIleGlnHisLeuAlaArg 353

Qy       1480 GAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGA 1539
              :::||||||   |||:::|||:::|||::::::||||||:::||||||
Db        354 AspPheLeu---SerAspTyrIlePheLeuSerValGlyArgValGlySerThrSerGlu 372

Qy       1540 GATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAA 1599
              ::::::   |||   ||||||   |||               |||:::      |||::::::
Db        373 AsnIleThrGlnLysValLeuTyrValGluAsnGlnAspLysLysSerAlaLeuLeuAsp 392

Qy       1600 ATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCA 1659
              :::|||                |||      |||::::::||||||||||||||||:::    |||
Db        393 LeuLeuSerAlaSerThrAspGlyLeuThrLeuIlePheValGluThrLysArgMetAla 412

Qy       1660 GATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGAT 1719
              |||         ||||||      :::            |||:::||||||||||||
Db        413 AspGlnLeuThrAspPheLeuIleMetGlnAsnPheArgAlaThrAlaIleHisGlyAsp 432

Qy       1720 CGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTT 1779
              |||   |||   |||||||||:::||||||      ||||||   |||            :::
Db        433 ArgThrGlnSerGluArgGluArgAlaLeuAlaAlaPheArgSerGlyAlaAlaThrLeu 452

Qy       1780 CTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATC 1839
              ||||||||||||:::||||||||||||||||||||||||   ||||||   |||||||||
Db        453 LeuValAlaThrAlaValAlaAlaArgGlyLeuAspIleProAsnValThrHisValIle 472

Qy       1840 AATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGT 1899
              |||:::||||||||||||      :::|||:::|||||||||||||||||||||||||||

Db        473 AsnTyrAspLeuProSerAspValAspAspTyrValHisArgIleGlyArgThrGlyArg 492

Qy       1900 TGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCA 1959
                 ||||||||||||   |||    :::||||||:::    |||:::    |||    :::
Db        493 AlaGlyAsnThrGlyLeuAlaThrAlaPhePheAsnSerGluAsnSerAsn---IleVal 511

Qy       1960 CAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAA 2019
              :::   |||   ::::::||||||:::|||   |||:::||||||::::::|||::::::
Db        512 LysGlyLeuHisGluIleLeuThrGluAlaAsnGlnGluValProSerPheLeuLysAsp 531

Qy       2020 ATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCA 2079
                          |||             ||||||                   |||
Db        532 AlaMetMetSerAla------ProGlySerArgSerAsnSerArgArgGlyGlyPheGly 549

Qy       2080 TCAGTTGATACCAGAAAGGGCAAGAGCACT------TTGAACACAGCTGGGTTTTCTTCT 2133
                    :::   |||         :::               :::         |||:::   |||
Db        550 ArgAsnAsnAsnArgAspTyrArgLysAlaGlyGlyAlaSerAlaGlyGlyTrpGlySer 569

Qy       2134 TCACGAGCT 2142
              ||||||:::
Db        570 SerArgSer 572
```

RESULT 11
T43543
probable ATP-dependent RNA helicase [similarity] - fission yeast (Schizosaccharomyces pom
N;Alternate names: suppressor of uncontrolled mitosis
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 19-Jan-2001
C;Accession: T43543; T41132; T43658; T50462; T43347
R;Forbes, K.L.C.; Enoch, T.
submitted to the EMBL Data Library, March 1998
A;Description: A suppressor of fission yeast checkpoint mutants encodes a putative ATP-de
A;Reference number: Z22562
A;Accession: T43543
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-636 <FOR>
A;Cross-references: EMBL:AF025536; PIDN:AAC04893.1
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Oliver, K.; Harris, D.
submitted to the EMBL Data Library, April 1998
A;Reference number: Z21971
A;Accession: T41132
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-636 <LYN>
A;Cross-references: EMBL:AL022598; PIDN:CAA18646.1; GSPDB:GN00068; SPDB:SPCC1795.11
A;Experimental source: strain 972h(-); cosmid c1795
R;Grallert, B.; Kearsey, S.E.; Lenhard, M.; Carlson, C.R.; Nurse, P.; Boye, E.; Labib, K
submitted to the EMBL Data Library, March 1999
A;Description: A putative RNA helicase reveals links between translation and the fission
A;Reference number: Z22607
A;Accession: T43658
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-636 <GRA>
A;Cross-references: EMBL:AJ237697; PIDN:CAB40192.1
R;Liu, H.Y.; Walworth, N.C.
submitted to the EMBL Data Library, August 1998
A;Description: A multi-copy suppressor of cdc2-r4, Dep1, is homologous to the Sacchromyce
A;Reference number: Z25074
A;Accession: T50462
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-636 <LIU>
A;Cross-references: EMBL:AF084222; PIDN:AAC34121.1
R;Kawamukai, M.
Biochim. Biophys. Acta 1446, 93-101, 1999
A;Title: Isolation of a novel gene, moc2, encoding a putative RNA helicase as a suppresso
A;Reference number: Z22361; MUID:99326140; PMID:10395922
A;Accession: T43347
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-40,'R',42-636 <KAW>

```
A;Cross-references: EMBL:AB012389; NID:g2992157; PIDN:BAA25324.1; PID:g2992158
C;Genetics:
A;Gene: sum3; ded1; SPCC1795.11; moc2
A;Map position: 3
C;Function:
A;Description: functions as suppressor of cdc2-r4, seems to have roles in both sexual di
A;Note: suppressor of sterile strains
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: nucleotide binding; P-loop
F;213-230/Region: nucleotide-binding motif A (P-loop)
F;332-337/Region: nucleotide-binding motif B
F;336-339/Region: DEAD motif

Alignment Scores:
Pred. No.:              1.46e-71       Length:         636
Score:                  1109.00        Matches:        271
Percent Similarity:     56.87%         Conservative:   89
Best Local Similarity:  42.81%         Mismatches:     185
Query Match:            28.05%         Indels:         88
DB:                     2              Gaps:           17

US-09-714-865-15 (1-2172) x T43543 (1-636)

Qy     382 TCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGA 441
Db      54 SerSerArgGlySerSerArgSerGlnThrProSerGluPheSerSerAsnTyr------ 71

Qy     442 GGTGGAAGAGGTAGTTTC---CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAAT 498
Db      72 GlyGlyArgArgGluTyrAsnArgGlyGlyHisTyrGlyGlyGlyGluGlyArgGlnAsn 91

Qy     499 AATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTGGTTCTAGAAGA 558
Db      92 Asn--------------------------TyrArgGlyGlyArgGluGly--------- 99

Qy     559 CCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT 618
Db     100 ------------GlyTyrSerAsnGlyGlyGlyTyrArgAsnAsnArgGlyPheGlyGln 115

Qy     619 GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT 678
Db     116 TrpArgAspGly---------------GlnHisValIle------GlyAlaArgAsnThr 128

Qy     679 TGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTAC 738
Db     129 LeuLeuGluArgGlnLeuPheGlyAlaValAlaAsp------GlyThrLysValSer--- 145

Qy     739 ATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATA 798
Db     146 ---------------------------------------------------ThrGlyIle 148

Qy     799 AACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCA 858
Db     149 AsnPheGluLysTyrAspAspIleProValGluValSerGlyGlyAspIleGluPro--- 167

Qy     859 ATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCT 918
Db     168 ValAsnGluPheThrSerProProLeuAsnSerHisLeuLeuGlnAsnIleLysLeuSer 187

Qy     919 GGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGA 978
Db     188 GlyTyrThrGlnProThrProValGlnLysAsnSerIleProIleValThrSerGlyArg 207

Qy     979 GATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
Db     208 AspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaGlyPheLeuPheProIle 227

Qy    1039 TTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG------------ 1086
Db     228 LeuSerLeuAlaPheAspLysGlyProAlaAlaValProValAspGlnAspAlaGlyMet 247

Qy    1087 ---------------TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTG 1131
Db     248 GlyTyrArgProArgLysAlaTyrProThrThrLeuIleLeuAlaProThrArgGluLeu 267

Qy    1132 GTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTT 1191
Db     268 ValCysGlnIleHisGluGluSerArgLysPheCysTyrArgSerTrpValArgProCys 287

Qy    1192 GTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAAT 1251
Db     288 AlaValTyrGlyGlyAlaAspIleArgAlaGlnIleArgGlnIleAspGlnGlyCysAsp 307

Qy    1252 ATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTC 1311
Db     308 LeuLeuSerAlaThrProGlyArgLeuValAspLeuIleAspArgGlyArgIleSerLeu 327

Qy    1312 AAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGT 1371
Db     328 AlaAsnIleLysPheLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlu 347

Qy    1372 CCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACC 1431
Db     348 ProGlnIleArgHisIleValGluGlyAlaAspMetThrSerValGluGluArgGlnThr 367

Qy    1432 CTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAG 1491
Db     368 LeuMetPheSerAlaThrPheProArgAspIleGlnLeuLeuAlaArgAspPheLeuLys 387

Qy    1492 TCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAG 1551
Db     388 ---AspTyrValPheLeuSerValGlyArgValGlySerThrSerGluAsnIleThrGln 406

Qy    1552 ACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611
Db     407 LysValValHisValGluAspSerGluLysArgSerTyrLeuLeuAspIleLeuHisThr 426

Qy    1612 ATAGGGGATGAA---AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACT 1668
Db     427 LeuProProGluGlyLeuThrLeuIlePheValGluThrLysArgMetAlaAspThrLeu 446

Qy    1669 GCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAG 1728
Db     447 ThrAspTyrLeuLeuAsnSerAsnPheProAlaThrSerIleHisGlyAspArgThrGln 466

Qy    1729 AGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCT 1788
Db     467 ArgGluArgGluArgAlaLeuGluLeuPheArgSerGlyArgThrSerIleMetValAla 486

Qy    1789 ACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGAT 1848
Db     487 ThrAlaValAlaSerArgGlyLeuAspIleProAsnValThrHisValIleAsnTyrAsp 506

Qy    1849 CTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAAT 1908
Db     507 LeuProThrAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAlaGlyAsn 526

Qy    1909 ACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTA 1968
Db     527 ThrGlyGlnAlaValAlaPhePheAsn---ArgAsnAsnLysGlyIleAlaLysGluLeu 545

Qy    1969 GTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTT 2028
Db     546 IleGluLeuLeuGlnGluAlaAsnGlnGluCysProSerPheLeuIleAlaMetAlaArg 565

Qy    2029 AGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGAT 2088
Db     566 GluSer---------SerPheGlyGlyAsnGlyArgGlyGlyArgTyrSerGlyArgGly 582

Qy    2089 ACCAGAAAGGGC--------------------------AAGAGCACTTTGAACACAGCT 2121
Db     583 GlyArgGlyGlyAsnAlaTyrGlyAlaArgAspPheArgArgProThrAsnSerSerSer 602

Qy    2122 GGGTTTTCTTCT-------------------------------TCACGAGCTCCC--- 2145
Db     603 GlyTyrSerSerGlyProSerTyrSerGlyTyrGlyGlyPheGluSerArgThrProHis 622
```

```
Qy       2146 ---------------AATCCAGTAGATGATGAGTCATGG 2169
                              |||          :::||||||
Db        623 HisGlyAsnThrTyrAsnSerGlySerAlaGlnSerTrp 635

RESULT 12
S62003
probable ATP-dependent RNA helicase DBP1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: probable RNA helicase CA1; protein LPH8c; protein YPL119c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 19-Jan-2001
C;Accession: S62003; S16790; A34848
R;Schlenstedt, G.; Silver, P.A.
submitted to the EMBL Data Library, December 1995
A;Reference number: S61996
A;Accession: S62003
A;Molecule type: DNA
A;Residues: 1-617 <SCH>
A;Cross-references: EMBL:U43503; NID:g1163087; PIDN:AAB68243.1; PID:g1163095; MIPS:YPL11
R;Jamieson, D.J.; Beggs, J.D.
Mol. Microbiol. 5, 805-812, 1991
A;Title: A suppressor of yeast spp81/ded1 mutations encodes a very similar putative ATP-
A;Reference number: S16790; MUID:91312117; PMID:1857205
A;Accession: S16790
A;Molecule type: DNA
A;Residues: 1-42,'RS',45-47,'K',49-87,'R',89-114,'QK',116-617 <JAM>
A;Cross-references: EMBL:X55993; NID:g3640; PIDN:CAA39465.1; PID:g3641
A;Experimental source: strain DBY939
R;Chang, T.H.; Arenas, J.; Abelson, J.
Proc. Natl. Acad. Sci. U.S.A. 87, 1571-1575, 1990
A;Title: Identification of five putative yeast RNA helicase genes.
A;Reference number: A34848; MUID:90160368; PMID:2406722
A;Accession: A34848
A;Molecule type: DNA
A;Residues: 316-495,'I',497-500 <CHA>
C;Genetics:
A;Gene: SGD:DBP1
A;Cross-references: SGD:S0006040; MIPS:YPL119c
A;Map position: 16L
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;198-205/Region: nucleotide-binding motif A (P-loop)
F;314-319/Region: nucleotide-binding motif B
F;318-321/Region: DEAD motif

Alignment Scores:
Pred. No.:                1.71e-71     Length:         617
Score:                    1108.00      Matches:        252
Percent Similarity:       57.43%       Conservative:   92
Best Local Similarity:    42.07%       Mismatches:     185
Query Match:              28.02%       Indels:         70
DB:                       2            Gaps:           14

US-09-714-865-15 (1-2172) x S62003 (1-617)

Qy        409 AATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC--------- 459
              ||||||   |||     |||          ||||||:::  |||:::
Db         14 AsnAsnLysGluAsnGlyGly---------GlyGlyGlyLysSerSerTyrValProPro 30

Qy        460 ------------CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTA 507
                          ||||||   :::    |||        :::|||        :::
Db         31 HisLeuArgSerArgGly---LysProSerPheGluArgSerThrProLysGlnGlu--- 48

Qy        508 GACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTA 567
                    |||:::          |||||||||   |||      |||
Db         49 ------AspLysVal---------ThrGlyGlyAspPhePheArgArg------------ 59

Qy        568 AGTGGC------ACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAA 621
              :::|||      |||||||||
Db         60 AlaGlyArgGlnThrGlyAsn--------------------------------------- 66

Qy        622 CGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGG 681
              ||||||:::  |||   ::::::|||          |||:::      |||
Db         67 AsnGlyGlyPhePheGlyPheSerLysGluArgAsnGlyGlyThrSerAlaAsnTyrAsn 86

Qy        682 AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATA 741
              :::   :::      :::||||||     :::               :::|||
Db         87 ArgGlyGlySerSerAsnTyrLysSerSerGlyAsnArgTrpValAsnGlyLysHisIle 106

Qy        742 CCC---CCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA------------------ 780
              |||   |||         :::   |||      :::|||
Db        107 ProGlyProLysAsnAlaLysLeuGluAlaGluLeuPheGlyValHisAspAspProAsp 126

Qy        781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
              ::::::   :::||||||   ||||||   ||||||   |||   |||:::   ||||||
Db        127 TyrHisSerSerGlyIleLysPheAspAsnTyrAspAsnIleProValAspAlaSerGly 146

Qy        841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
                 |||   |||      ||||||   |||            |||   :::   |||
Db        147 LysAspValProGluProIleLeuAspPheSerSerProProLeuAspGluLeuLeuMet 166

Qy        901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
                 ||||||   |||   :::||||||   ||||||||||||||||||||||||
Db        167 GluAsnIleLysLeuAlaSerPheThrLysProThrProValGlnLysTyrSerIlePro 186

Qy        961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
              |||:::         |||||||||||||||||||||||||||||||||||||||||
Db        187 IleValThrLysGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrGly 206

Qy       1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACT------------ 1068
              ||||||   ||||:::            :::         |||   :::
Db        207 GlyPheLeuPheProLeuPheThrGluLeuPheArgSerGlyProSerProValProGlu 226

Qy       1069 ------GCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACT 1122
                    |||   :::           |||         ::::::::|||||||||
Db        227 LysAlaGlnSerPheTyrSerArgLysGlyTyrProSerAlaLeuValLeuAlaProThr 246

Qy       1123 CGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTA 1182
              |||||||||      ||||||:::   |||||||||||||||:::::::   :::   |||
Db        247 ArgGluLeuAlaThrGlnIlePheGluGluAlaArgLysPheThrTyrArgSerTrpVal 266

Qy       1183 AGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAA 1242
              |||      |||:::|||||||||      :::|||:::   :::|||::::::   :::
Db        267 ArgProCysValValTyrGlyGlyAlaProIleGlyAsnGlnMetArgGluValAspArg 286

Qy       1243 GGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAG 1302
              ||||||::::::|||   ||||||||||||||||||   |||::::::   :::   |||
Db        287 GlyCysAspLeuLeuValAlaThrProGlyArgLeuAsnAspLeuLeuGluArgGlyLys 306

Qy       1303 ATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATG 1362
              :::   |||   |||||||||||||||||||||||||||||||||||||||||||||
Db        307 ValSerLeuAlaAsnIleLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMet 326

Qy       1363 GGTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAG 1422
              ||||||   |||:::::::::   ::::::         |||||||||   |||
Db        327 GlyPheGluProGlnIleArgHisIleValGluGluCysAspMetProSerValGluAsn 346

Qy       1423 CGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAG 1482
              ||||||||||||||||||||||||||||||   :::||||||   ||||||   :::
Db        347 ArgGlnThrLeuMetPheSerAlaThrPheProValAspIleGlnHisLeuAlaArgAsp 366

Qy       1483 TTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGAT 1542
              ||||||      ||||||:::|||::::::||||||:::||||||         :::
Db        367 PheLeu---AspAsnTyrIlePheLeuSerValGlyArgValGlySerThrSerGluAsn 385

Qy       1543 GTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATT 1602
              :::   |||   :::|||   |||         |||:::      |||::::::::
Db        386 IleThrGlnArgIleLeuTyrValAspAspMetAspLysLysSerAlaLeuLeuAspLeu 405

Qy       1603 CTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGAT 1662
              |||                  |||::::::|||||||||||||||:::   ||||||
Db        406 LeuSerAlaGluHisLysGlyLeuThrLeuIlePheValGluThrLysArgMetAlaAsp 425

Qy       1663 TTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGG 1722
                       ||||||      :::            |||:::||||||||||||||
Db        426 GlnLeuThrAspPheLeuIleMetGlnAsnPheLysAlaThrAlaIleHisGlyAspArg 445
```

```
Qy       1723 GAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTT 1782
                 |||   ||||||||:::||||||      |||:::              :::|||
Db        446 ThrGlnAlaGluArgGluArgAlaLeuSerAlaPheLysAlaAsnValAlaAspIleLeu 465

Qy       1783 GTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAAT 1842
              |||||||||:::|||||||||||||||||||||||||   ||||||   |||||||||||
Db        466 ValAlaThrAlaValAlaAlaArgGlyLeuAspIleProAsnValThrHisValIleAsn 485

Qy       1843 TTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGT 1902
              :::|||||||||||   ||||||:::||||||||||||||||||||||||||||||
Db        486 TyrAspLeuProSerAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAla 505

Qy       1903 GGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAG 1962
              |||||||||||   |||   |||||||||:::      :::   |||   :::   :::
Db        506 GlyAsnThrGlyValAlaThrSerPhePheAsnSerAsnAsnGlnAsn---IleValLys 524

Qy       1963 CCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATT 2022
                 |||:::::::::|||   :::|||   |||:::||||||   :::|||   ::::::
Db        525 GlyLeuMetGluIleLeuAsnGluAlaAsnGlnGluValProThrPheLeuSerAspLeu 544

Qy       2023 GCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCA 2082
                                                                       |||
Db        545 ---------------------------------------------------------Ser 545

Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
                          :::|||      |||            |||||   :::||||||
Db        546 ArgGlnAsnSerArgGlyGlyArgThrArgGlyGlyGlyGlyPhePheAsnSerArg 564
```

RESULT 13
H84854
probable ATP-dependent RNA helicase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: H84854
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84854
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-633 <STO>
A;Cross-references: GB:AE002093; NID:g4559339; PIDN:AAD23001.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g42520
A;Map position: 2
C;Superfamily: ATP-dependent RNA helicase DBP1

```
Alignment Scores:
Pred. No.:                6.46e-71      Length:          633
Score:                    1100.00       Matches:         260
Percent Similarity:       55.71%        Conservative:    96
Best Local Similarity:    40.69%        Mismatches:      191
Query Match:              27.82%        Indels:          92
DB:                       2             Gaps:            16
```

US-09-714-865-15 (1-2172) x H84854 (1-633)

```
Qy        307 GATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACA 366
              |||   :::::::::|||               :::   |||            |||
Db         10 AspSerGluAsnThrGlySerGlySerSerAsnGlnAsnSerHisProSerArgProAla 29

Qy        367 -----------------CGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGA 408
                               |||||||||         :::
Db         30 TyrValProProHisLeuArgAsnArgProAlaAla------------------------ 41

Qy        409 AATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGT--- 465
                    ||||||      |||              |||      :::   |||
Db         42 ------SerGluProValAlaProLeuProAlaAsnAspArgValGlyTyrGlyGlyPro 59

Qy        466 -----------------TGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTA 507
                               ||||||   |||:::|||
Db         60 ProSerGlySerArgTrpAlaProGlyGlySerGlyValGly------------------ 73

Qy        508 GACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTGGTTCTAGAAGACCAGTATTA 567
                                                                        :::
Db         74 ---------------------------------------------------------Val 74

Qy        568 AGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGT 627
                 |||   |||                           |||   |||||||||||   |||
Db         75 GlyGlyGlyGlyGlyTyrArgAlaAspAlaGlyArgProGlySerGlySerGlyTyrGly 94

Qy        628 GGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT--------- 678
              |||   :::|||                           |||   |||   |||:::
Db         95 Gly---ArgGly--------------------GlyGlyGlyTrpAsnAsnArgSerGly 106

Qy        679 ---TGG---------AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCA 726
                 |||         :::   |||         |||:::   |||::::::
Db        107 GlyTrpAspArgArgGluArgGluValAsnProPheGluAsnAspAspSerGlu------ 124

Qy        727 AAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTAT 786
                             |||   |||         ::::::|||:::
Db        125 ---------------ProGluProAlaPheThrGluGlnAspAsn--------------- 134

Qy        787 CAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGAT 846
                 |||   ||||||||||||   |||:::   |||   :::|||   ||||||   :::
Db        135 ---ThrValIleAsnPheAspAlaTyrGluAspIleProIleGluThrSerGlyAspAsn 153

Qy        847 GCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAAC 906
                 ||||||   :::   ||||||   |||   :::|||   :::   ||||||   |||
Db        154 ValProProProValAsnThrPheAlaGluIleAspLeuGlyGluAlaLeuAsnLeuAsn 173

Qy        907 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 966
              |||   :::         |||   |||   ||||||||||||:::::::::|||||||||:::
Db        174 IleArgArgCysLysTyrValLysProThrProValGlnArgHisAlaIleProIleLeu 193

Qy        967 CTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 LeuGluGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPhe 213

Qy       1027 CTCCTACCAATTTTGGCTCATATGATGCATGAT---GGAATAACTGCCAGTCGTTTTAAA 1083
                 ||||||::::::   :::|||   |||         :::         |||
Db        214 CysPheProIleIleSerGlyIleMetLysAspGlnHisValGlnArgProArgGlySer 233

Qy       1084 GAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATT 1143
                       |||      :::|||::::::||||||||||||||||   :::||||||
Db        234 ArgThrValTyrProLeuAlaValIleLeuSerProThrArgGluLeuAlaSerGlnIle 253

Qy       1144 TATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGG 1203
              :::   ||||||:::|||||||||:::   |||   |||:::   ||||||   ||||||
Db        254 HisAspGluAlaLysLysPheSerTyrGlnThrGlyValLysValValValAlaTyrGly 273

Qy       1204 GGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCT 1263
              ||||||   :::         :::|||::::::   :::|||   :::||||||   |||
Db        274 GlyThrProIleAsnGlnGlnLeuArgGluLeuGluArgGlyValAspIleLeuValAla 293

Qy       1264 ACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAA 1323
              |||||||||||||||   |||::::::   :::   ::::::   ::::::   |||:::
Db        294 ThrProGlyArgLeuAsnAspLeuLeuGluArgAlaArgValSerMetGlnMetIleArg 313

Qy       1324 TACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAG 1383
              :::|||   |||||||||||||||||||||||||||||||||||   |||:::::::::
Db        314 PheLeuAlaLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArg 333

Qy       1384 AAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGT 1443
              |||::::::               ||||||   :::      ||||||||||||:::||||||
Db        334 LysIleValGluGlnMetAspMetProProArgGlyValArgGlnThrLeuLeuPheSer 353

Qy       1444 GCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTG 1503
              ||||||||||||   |||||||||||||||||||:::||||||      :::||||||:::
Db        354 AlaThrPheProArgGluIleGlnArgLeuAlaAlaAspPheLeu---AlaAsnTyrIle 372

Qy       1504 TTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAA 1563
```

```
                  |||:::||||||||:::||||||   :::           :::   |||   |||
Db         373 PheLeuAlaValGlyArgValGlySerSerThrAspLeuIleValGlnArgValGluPhe 392

Qy        1564 GTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTG---------CGAAACATA 1614
                |||          ||||||      |||:::::::::|||           |||
Db         393 ValLeuAspSerAspLysArgSerHisLeuMetAspLeuLeuHisAlaGlnArgGluAsn 412

Qy        1615 GGG------------GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGAT 1662
               |||                      |||:::||||||||||||||||:::   ||||||
Db         413 GlyIleGlnGlyLysGlnAlaLeuThrLeuValPheValGluThrLysArgGlyAlaAsp 432

Qy        1663 TTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGG 1722
                        :::||||||                 ||||||||||||||||||
Db         433 SerLeuGluAsnTrpLeuCysIleAsnGlyPheProAlaThrSerIleHisGlyAspArg 452

Qy        1723 GAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTT 1782
                  |||:::|||||||||   ||||||      |||:::   |||:::   |||:::|||
Db         453 ThrGlnGlnGluArgGluValAlaLeuLysAlaPheLysSerGlyArgThrProIleLeu 472

Qy        1783 GTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAAT 1842
               |||||||||   |||||||||||||||||||||||||   :::|||   ||||||:::|||
Db         473 ValAlaThrAspValAlaAlaArgGlyLeuAspIleProHisValAlaHisValValAsn 492

Qy        1843 TTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGT 1902
               |||||||||:::   ||||||:::||||||||||||||||||||||||||||||
Db         493 PheAspLeuProAsnAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAla 512

Qy        1903 GGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAG 1962
               |||   :::|||   |||   :::||||||:::   :::   :::      ||||||:::
Db         513 GlyLysSerGlyLeuAlaThrAlaPhePheAsn---AspGlyAsnThrSerLeuAlaArg 531

Qy        1963 CCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATT 2022
               ||||||   :::::::::   :::|||   |||:::||||||   ||||||
Db         532 ProLeuAlaGluLeuMetGlnGluAlaAsnGlnGluValProGluTrpLeuThrArgTyr 551

Qy        2023 GCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCA 2082
               |||         :::            |||               |||      |||
Db         552 AlaSerArgSerSerPheGlyGlyGlyLysAsnArgArgSerGlyGlyArgPheGlyGly 571

Qy        2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
                  |||   |||:::      |||                  |||   |||   |||
Db         572 ArgAspPheArgArgGluGlySer------------------PheGlySerGlyArg 584
```

RESULT 14
T45677
ATP-dependent RNA helicase-like protein - Arabidopsis thaliana
N;Alternate names: protein F14P22.160
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
C;Accession: T45677
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23011
A;Accession: T45677
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-646 <DAN>
A;Cross-references: EMBL:AL137082
A;Experimental source: cultivar Columbia; BAC clone F14P22
C;Genetics:
A;Map position: 3
A;Introns: 239/3; 267/3; 348/3; 404/3; 442/3
A;Note: F14P22.160
C;Superfamily: ATP-dependent RNA helicase DBP1

```
Alignment Scores:
Pred. No.:               1.37e-70      Length:        646
Score:                   1095.50       Matches:       258
Percent Similarity:      57.33%        Conservative:  90
Best Local Similarity:   42.50%        Mismatches:    194
Query Match:             27.71%        Indels:        65
DB:                      2             Gaps:          14
```

```
US-09-714-865-15 (1-2172) x T45677 (1-646)

Qy         337 TCTAGTAATGACTGCGAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGC--- 393
               |||||||||         |||                  |||      |||      |||
Db           2 SerSerAsnSerTrpAlaAspValSerGluSerGluArgAlaProSerGlyGlyGlyTrp 21

Qy         394 GGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGT 453
               ||||||         :::   :::                  |||:::   |||
Db          22 GlyTyrSerArgProSerArgThrAsnTyrValProProHisLeuArgSer--------- 38

Qy         454 AGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCA---AATAATGACTTAGAC 510
                     |||               |||         ||||||   |||||||||
Db          39 ------ArgThrProSerSerGluPheValAlaProSerProGlyAsnAsnAsp------ 54

Qy         511 CCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGT 570
                                       ||||||   |||                     :::
Db          55 ---------------------ArgGlyGlyTyrGlyGly------------------Ala 61

Qy         571 GGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGT 630
                  :::|||   |||         |||         |||   ||||||      ||||||
Db          62 AsnSerGlyTyrGlyGlyArgGlyGlnGlyTyrGlyGlyArgGlySerGlyTyrGlyGly 81

Qy         631 TACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAA 690
                                             |||            |||   :::
Db          82 ---------------------------ArgGlyGlyProValGlyGlyTrpAsnAlaArg 92

Qy         691 GCAGAAGGAGGAGAAAGTAGTGATACTCAA------------------------------ 720
               :::   |||   :::      ||||||:::
Db          93 SerGlyGlyTrpAspArgArgAspThrGluThrAsnProPheGlyAsnAspGlyAsnAla 112

Qy         721 GGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCA 780
               |||   |||                     |||   |||:::::::::
Db         113 AspProAlaValAsn---------------------GluGlnGluAsnThrVal------ 123

Qy         781 CATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGA 840
                              |||||||||:::   |||:::   |||   :::|||   ||||||
Db         124 ---------------IleAsnPheGluAlaTyrGluAspIleProIleGluThrSerGly 138

Qy         841 CATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAAT 900
                  :::   ||||||   :::   ||||||   |||   :::|||   :::   ||||||
Db         139 AspAsnValProProProValAsnThrPheAlaGluIleAspLeuGlyGluAlaLeuAsn 158

Qy         901 AACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCT 960
                  ||||||   :::      |||   |||   ||||||||||||:::   :::||||||
Db         159 LeuAsnIleGlnArgCysLysTyrValLysProThrProValGlnArgAsnAlaIlePro 178

Qy         961 ATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCG 1020
               |||:::   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         179 IleLeuAlaAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 198

Qy        1021 GCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGAT---GGAATAACTGCCAGTCGT 1077
               ||||||      ||||||::::::   :::|||   |||      |||         |||
Db         199 AlaPheCysPheProIleIleSerGlyIleMetLysAspGlnHisIleGluArgProArg 218

Qy        1078 TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAAC 1137
                                    |||      :::|||::::::||||||||||||||
Db         219 GlyValArgGlyValTyrProLeuAlaValIleLeuSerProThrArgGluLeuAlaCys 238

Qy        1138 CAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATA 1197
               ||||||:::   |||||||||||||||||:::   |||   |||:::   ||||||
Db         239 GlnIleHisAspGluAlaArgLysPheSerTyrGlnThrGlyValLysValValValAla 258

Qy        1198 TATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTA 1257
               ||||||||||||   :::         ||||||::::::   :::|||   :::||||||
Db         259 TyrGlyGlyThrProValAsnGlnGlnIleArgGluLeuGluArgGlyValAspIleLeu 278

Qy        1258 TGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAG 1317
                  ||||||||||||||||||   |||::::::   :::   ::::::   |||:::
Db         279 ValAlaThrProGlyArgLeuAsnAspLeuLeuGluArgGlyArgValSerLeuGlnMet 298

Qy        1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAA 1377
               :::::::::|||   ||||||||||||||||||||||||||||||||||||   |||:::
Db         299 ValArgPheLeuAlaLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGln 318
```

```
Qy       1378 ATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATG 1437
Db        319 IleArgLysIleValGlnGlnMetAspMetProProProGlyValArgGlnThrMetLeu 338

Qy       1438 TTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAAT 1497
Db        339 PheSerAlaThrPheProArgGluIleGlnArgLeuAlaSerAspPheLeu---SerAsn 357

Qy       1498 TATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT 1557
Db        358 TyrIlePheLeuAlaValGlyArgValGlySerSerThrAspLeuIleValGlnArgVal 377

Qy       1558 CTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTG---------CGA 1608
Db        378 GluPheValHisAspSerAspLysArgSerHisLeuMetAspLeuLeuHisAlaGlnArg 397

Qy       1609 AACATAGGGGATGAA------------AGAACTATGGTCTTTGTTGAAACTAAGAAAAAA 1656
Db        398 GluAsnGlyAsnGlnGlyLysGlnAlaLeuThrLeuValPheValGluThrLysLysGly 417

Qy       1657 GCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGT 1716
Db        418 AlaAspSerLeuGluAsnTrpLeuCysIleAsnGlyPheProAlaThrThrIleHisGly 437

Qy       1717 GATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCA 1776
Db        438 AspArgSerGlnGlnGluArgGluValAlaLeuArgSerPheLysThrGlyArgThrPro 457

Qy       1777 GTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTT 1836
Db        458 IleLeuValAlaThrAspValAlaAlaArgGlyLeuAspIleProHisValAlaHisVal 477

Qy       1837 ATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGT 1896
Db        478 ValAsnPheAspLeuProAsnAspIleAspAspTyrValHisArgIleGlyArgThrGly 497

Qy       1897 CGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTA 1956
Db        498 ArgAlaGlyAsnSerGlyLeuAlaThrAlaPhePheAsn---AspAsnAsnThrThrMet 516

Qy       1957 GCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAA 2016
Db        517 AlaLysProLeuAlaGluLeuMetGlnGluAlaAsnGlnGluValProAspTrpLeuThr 536

Qy       2017 GAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTT 2076
Db        537 ArgTyrAlaSerArgAlaSerPheGlyGlyGlyLysAsnArgArgSerGlyGlyArgPhe 556

Qy       2077 GCATCAGTTGATACCAGAAAG 2097
Db        557 GlyGlyArgAspPheArgArg 563
```

RESULT 15
A48686
probable RNA helicase glh-1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Aug-2001
C;Accession: A48686
R;Roussell, D.L.; Bennett, K.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 9300-9304, 1993
A;Title: glh-1, a germ-line putative RNA helicase from Caenorhabditis, has four zinc fin
A;Reference number: A48686; MUID:94022363; PMID:8415696
A;Accession: A48686
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-707 <ROU>
A;Cross-references: GB:L19948
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;329-336/Region: nucleotide-binding motif A (P-loop)
F;439-444/Region: nucleotide-binding motif B
F;443-446/Region: DEAD motif

```
Alignment Scores:
Pred. No.:              1.01e-69       Length:          707
Score:                  1083.50        Matches:         269
Percent Similarity:     49.27%         Conservative:    100
Best Local Similarity:  35.91%         Mismatches:      269
Query Match:            27.40%         Indels:          111
DB:                     2              Gaps:            16

US-09-714-865-15 (1-2172) x A48686 (1-707)

Qy        130 ATGGATGATGGACCTTCTCGAAGAGATCATTTC------ATGAAAAGTGGATTTGCCTCT 183
Db          1 MetSerAspGlyTrpSerAspSerGluSerAlaAlaLysAlaLysThrGlyPheGlySer 20

Qy        184 GGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGT------------AATAAGCGAGAT 231
Db         21 GlyGlyGlyPheGlyGlyGlyAsnAsnGlyGlySerGlyPheGlyGlyGlyLysAsnGly 40

Qy        232 AATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGTTTGGAAACAGAGGTTTTTCA 291
Db         41 GlyThrGlyPheGlyGlyGlyAsnThrGlyGlySerGlyPheGlyGlyGlyAsnThrGly 60

Qy        292 AACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC 351
Db         61 GlySerGlyPheGlyGlyGlyLysThrGlyGlySerGlyPheGlyGlyGlyAsnThrCys 80

Qy        352 GAAGATAATCCAACACGGAACAGAGGGTTTTCCAAG---------AGAGGCGGCTATCGA 402
Db         81 GlySerPheGlyGlyGlyAsnSerGlyPheGlyGluGlyGlyHisGlyGlyGlyGluArg 100

Qy        403 GATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTCCGA 462
Db        101 AsnAsnAsnCysPheAsnCysGlnGlnPro---------GlyHisArgSerSer------ 115

Qy        463 GGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGT 522
Db        116 ------------------AspCysProGluProArgLysGluArgGluProArgValCys 129

Qy        523 ATGCAG------------------------------------------------------ 528
Db        130 TyrAsnCysGlnGlnProGlyHisThrSerArgGluCysThrGluGluArgLysProArg 149

Qy        529 ------CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAAT 582
Db        150 GluGlyArgThrGlyGlyPheGlyGlyGlyAlaGlyPheGlyAsnAsnGlyGlyAsnAsp 169

Qy        583 GGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTA 642
Db        170 GlyPheGlyGlyAspGlyGlyPheGlyGlyGlyGluGluArgGlyProMetLysCysPhe 189

Qy        643 AATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAG------------------ 684
Db        190 Asn---------CysLysGlyGluGlyHisArgSerAlaGluCysProGluProProArg 206

Qy        685 ------------------------------------------------TCAGAAGCAGAA 696
Db        207 GlyCysPheAsnCysGlyGluGlnGlyHisArgSerIleGluCysProAsnProAlaLys 226

Qy        697 GGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCT 756
Db        227 ProArgGluGlyValAspGlyGluGlyProLysAlaThrTyrValProValGluAspAsn 246

Qy        757 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGAC 816
Db        247 MetGluAspValPheAsnMetGlnLysIleSerGluGlyLeuMetPheAsnLysPhePhe 266

Qy        817 ACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATT---------CTGACT 867
Db        267 AspAlaGluValLysLeuThrSerSerGluLysThrValGlyIleLysProCysLysThr 286

Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
```

```
Db        287 PheAlaGluAlaAsnLeuThrGluThrMetGlnLysAsnValAlaHisAlaGlyTyrSer 306

Qy        928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
              |||   ||||||:::|||:::|||::::::|||::::::        |||    |||:::|||
Db        307 LysThrThrProIleGlnGlnTyrAlaLeuProLeuValHisGlnGlyTyrAspIleMet 326

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCAT 1047
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||:::
Db        327 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuSerIleMetThrArg 346

Qy       1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
              ::::::    |||              :::      :::          |||    ||||||
Db        347 LeuIleAspAspAsnAsnLeuAsnThrAlaGlyGluGlyGlyCysTyrProArgCysIle 366

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
              |||:::   ||||||||||||||    :::|||||||||   |||    |||||||||:::
Db        367 IleLeuThrProThrArgGluLeuAlaAspGlnIleTyrAsnGluGlyArgLysPheAla 386

Qy       1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
              :::    |||     :::                :::|||||||||     :::|||::::::
Db        387 TyrGlnThrMetMetGluIleLysProValTyrGlyGlyLeuAlaValGlyTyrAsnLys 406

Qy       1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
                 ||||||   :::|||      |||:::         |||   ||||||:::
Db        407 GlyGlnIleGluLysGlyAlaThrIleIleValGlyThrValGlyArgIleLysHisPhe 426

Qy       1288 ATAGGCAAAGAAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGAT 1347
                    :::        |||   |||    :::    ::::::     ||||||||||||||||||
Db        427 CysGluGluGlyThrIleLysLeuAspLysCysArgPhePheValLeuAspGluAlaAsp 446

Qy       1348 CGCATGTTGGAT---ATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
              ||||||:::|||   ||||||||||||    :::::::::   :::::::::
Db        447 ArgMetIleAspAlaMetGlyPheGlyThrAspIleGluThrIleValAsnTyrAspSer 466

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
              ||||||   ||||||   |||||||||||||||||||||||||||||||||:::    :::
Db        467 MetProArgLysGluAsnArgGlnThrLeuMetPheSerAlaThrPheProAspSerVal 486

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
              |||       |||         ||||||:::    ||||||:::   :::|||:::    ::::::
Db        487 GlnGluAlaAlaArgAlaPheLeuArgGluAsnTyrValMetIleAlaIleAspLysIle 506

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
              |||   |||    :::   |||   |||               :::       :::       |||:::
Db        507 GlyAlaAlaAsnLysCysValLeuGlnGluPheGluArgCysGluArgSerGluLysLys 526

Qy       1585 GAAAAGCTCGTTGAAATTCTG-----------------------------------CGA 1608
              :::||||||:::|||:::|||
Db        527 AspLysLeuLeuGluLeuLeuGlyIleAspIleAspSerTyrThrThrGluLysSerAla 546

Qy       1609 AACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACT 1668
                 :::      ::::::||||||||||||||           :::        ||||||
Db        547 GluValTyrThrLysLysThrMetValPheValSerGlnArgAlaMetAlaAspIleLeu 566

Qy       1669 GCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAG 1728
              |||:::    |||              ::::::             :::|||||||||   |||||||||
Db        567 AlaSerIleLeuSerSerAlaGlnValProAlaIleThrIleHisGlyAlaArgGluGln 586

Qy       1729 AGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCT 1788
              |||||||||     :::||||||          ||||||   |||       |||||||||:::|||
Db        587 ArgGluArgSerGluAlaLeuArgGlnPheArgAsnGlySerLysProValLeuIleAla 606

Qy       1789 ACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGAT 1848
              |||:::||||||   ||||||||||||||:::      |||    ||||||||||||:::|||
Db        607 ThrAlaValAlaGluArgGlyLeuAspIleLysGlyValAspHisValIleAsnTyrAsp 626

Qy       1849 CTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAAT 1908
              :::|||      ||||||:::|||:::|||||||||||||||||||||||||      ||||||
Db        627 MetProAspAsnIleAspAspTyrIleHisArgIleGlyArgThrGlyArgValGlyAsn 646

Qy       1909 ACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTA 1968
              :::|||||||||     ||||||          :::       |||       |||          |||
Db        647 SerGlyArgAlaThrSerPhe------IleSerGluAspCysSerLeuLeuSerGluLeu 664

Qy       1969 GTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTT 2028
              |||   ||||||   ||||||||||||   ||||||   |||::::::
Db        665 ValGlyValLeuAlaAspAlaGlnGlnIleValProAspTrpMetGln------------ 680

Qy       2029 AGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGAT 2088
                                      |||:::      ||||||      ||||||
Db        681 ------------------------GlyAlaAlaGlyGlyAsnTyrGlyAlaSer------ 690

Qy       2089 ACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAAT 2148
                                                  ||||||   ||||||       |||
Db        691 ---------------------------------GlyPheGlySerSer---ValProThr 698

Qy       2149 ------CCAGTAGATGATGAGTCATGG 2169
                    |||   |||:::|||   |||
Db        699 GlnValProGlnAspGluGluGlyTrp 707
```

Search completed: June 10, 2003, 16:44:39
Job time : 115.5 secs

```
OM nucleic - protein search, using frame_plus_n2p model

Run on:         June 10, 2003, 16:42:51 ; Search time 156.5 Seconds
                                          (without alignments)
                                          2865.660 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  3954
Sequence:       1 atgggggatgaagattggga.........tagatgatgagtcatgggat 2172

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters:      784170

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111800_26140/app_query.fasta_1.2311
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09714865_@CGN_1_1_59_@runat_05062003_111800_26140
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_AA:*
                1:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                10:  /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                12:  /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
                13:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                14:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	796	20.1	575	9	US-10-108-605-55	Sequence 55, Appl
2	761	19.2	500	9	US-10-228-897-13	Sequence 13, Appl
3	709.5	17.9	648	10	US-09-923-831-43	Sequence 43, Appl
4	704	17.8	1261	12	US-10-147-268-2	Sequence 2, Appli
5	681.5	17.2	709	9	US-09-736-457-335	Sequence 335, App
6	681.5	17.2	709	9	US-09-902-941-335	Sequence 335, App
7	681.5	17.2	709	9	US-09-849-626-335	Sequence 335, App
8	681.5	17.2	709	9	US-10-017-754-335	Sequence 335, App
9	622	15.7	630	9	US-10-195-117-2	Sequence 2, Appli
10	621	15.7	524	10	US-09-815-242-13491	Sequence 13491, A
11	618.5	15.6	732	9	US-09-738-626-4778	Sequence 4778, Ap
12	614.5	15.5	492	10	US-09-815-242-11308	Sequence 11308, A
13	614.5	15.5	492	10	US-09-732-091-2	Sequence 2, Appli
14	614.5	15.5	504	10	US-09-732-091-42	Sequence 42, Appl
15	601.5	15.2	624	10	US-09-963-790A-2	Sequence 2, Appli
16	599.5	15.2	536	10	US-09-815-242-10774	Sequence 10774, A
17	587	14.8	495	9	US-09-971-536-52	Sequence 52, Appl
18	578	14.6	646	10	US-09-815-242-13769	Sequence 13769, A
19	575	14.5	646	10	US-09-815-242-10311	Sequence 10311, A
20	567	14.3	424	9	US-09-738-626-4359	Sequence 4359, Ap
21	562.5	14.2	613	10	US-09-815-242-10990	Sequence 10990, A
22	561.5	14.2	510	10	US-09-815-242-5256	Sequence 5256, Ap
23	561.5	14.2	526	9	US-10-084-205-74	Sequence 74, Appl
24	561.5	14.2	526	10	US-09-925-637-74	Sequence 74, Appl
25	552.5	14.0	487	10	US-09-815-242-12397	Sequence 12397, A
26	552.5	14.0	487	10	US-09-815-242-12991	Sequence 12991, A
27	552	14.0	567	10	US-09-815-242-11885	Sequence 11885, A
28	545.5	13.8	764	10	US-09-902-627-1	Sequence 1, Appli
29	542	13.7	106	9	US-09-764-891-3857	Sequence 3857, Ap
30	523.5	13.2	459	10	US-09-925-301-1027	Sequence 1027, Ap
31	517	13.1	455	9	US-09-909-567B-48	Sequence 48, Appl
32	515.5	13.0	421	10	US-09-815-242-11703	Sequence 11703, A
33	505.5	12.8	411	10	US-09-755-325-8	Sequence 8, Appli
34	504.5	12.8	415	10	US-09-755-325-2	Sequence 2, Appli
35	503.5	12.7	411	10	US-09-755-325-6	Sequence 6, Appli
36	502.5	12.7	343	10	US-09-925-301-1439	Sequence 1439, Ap
37	491.5	12.4	448	9	US-09-738-626-5103	Sequence 5103, Ap
38	486.5	12.3	447	10	US-09-815-242-13412	Sequence 13412, A
39	486.5	12.3	447	10	US-09-815-242-13590	Sequence 13590, A
40	485	12.3	403	9	US-09-991-496-10	Sequence 10, Appl
41	485	12.3	403	10	US-09-874-923-10	Sequence 10, Appl
42	472.5	11.9	452	10	US-09-815-242-10493	Sequence 10493, A
43	468.5	11.8	406	9	US-10-228-897-12	Sequence 12, Appl
44	467.5	11.8	506	10	US-09-801-368-90	Sequence 90, Appl
45	463.5	11.7	437	10	US-09-815-242-5233	Sequence 5233, Ap

ALIGNMENTS

```
RESULT 1
US-10-108-605-55
; Sequence 55, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
;  APPLICANT: Broadus, Julie
;  APPLICANT:  Stam, Lynn
;  APPLICANT:  Bachmann, Jane
;  APPLICANT:  Kamdar, Kim
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
;  TITLE OF INVENTION:  PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
;  FILE REFERENCE: 31133B
;  CURRENT APPLICATION NUMBER: US/10/108,605
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: US 09/761,142
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/176,418
;  PRIOR FILING DATE: 2000-01-14
;  NUMBER OF SEQ ID NOS: 361
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
;   LENGTH: 575
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-108-605-55

Alignment Scores:
Pred. No.:              2.1e-60        Length:         575
```

```
Score:                    796.00      Matches:        206
Percent Similarity:       51.18%      Conservative:    98
Best Local Similarity:    34.68%      Mismatches:     212
Query Match:              20.13%      Indels:          78
DB:                       9           Gaps:            17

US-09-714-865-15 (1-2172) x US-10-108-605-55 (1-575)

Qy      439 AGAGGTGGAAGAGGTAGTTTCCGAGGT---------TGCCGTGGAGGATTTGGTCTAGGA 489
Db       13 ArgGlyGlyArgGlyGlyAspArgGlyGlyAspAspArgArgGlyGlyGlyGlyGlyGly 32

Qy      490 AGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGT 549
Db       33 Asn-----------------------------------ArgPheGlyGly--------- 37

Qy      550 TCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGC 609
Db       38 --------------------GlyGlyGlyGlyGlyAspTyrHisGlyIleArgAsnGly 50

Qy      610 AGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGA 669
Db       51 AlaValGluLysArgArgAspAspArgGlyGlyGlyAsnArgPheGlyGlyGlyGlyGly 70

Qy      670 AAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAA 729
Db       71 PheGlyAsp---ArgArgGlyGlyGlyGlyGlyGlySerGlnAspLeu----------- 85

Qy      730 GTGACCTACATACCCCCTCCTCCACCTGAGGAT---------------------GAGGAC 768
Db       86 ---------------ProMetArgProValAspPheSerAsnLeuAlaProPheLysLys 100

Qy      769 TCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTT--- 825
Db      101 AsnPheTyrGlnGluHisProAsnValAlaAsnArgSerProTyrGluValGlnArgTyr 120

Qy      826 ---------------GTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTT 870
Db      121 ArgGluGluGlnGluIleThrValArgGly---GlnValProAsnProIleGlnAspPhe 139

Qy      871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
Db      140 SerGluValHisLeuProAspTyrValMetLysGluIleArgArgGlnGlyTyrLysAla 159

Qy      931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
Db      160 ProThrAlaIleGlnAlaGlnGlyTrpProIleAlaMetSerGlySerAsnPheValGly 179

Qy      991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG 1050
Db      180 IleAlaLysThrGlySerGlyLysThrLeuGlyTyrIleLeuProAlaIleValHisIle 199

Qy     1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA---------CCAGAG 1101
Db      200 AsnAsnGln--------------------GlnProLeuGlnArgGlyAspGlyProIle 212

Qy     1102 TGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAA 1161
Db      213 AlaLeuValLeuAlaProThrArgGluLeuAlaGlnGlnIleGlnGlnValAlaThrGlu 232

Qy     1162 TTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGGGACAT 1221
Db      233 PheGlySerSerSerTyrValArgAsnThrCysValPheGlyGlyAlaProLysGlyGly 252

Qy     1222 TCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATG 1281
Db      253 GlnMetArgAspLeuGlnArgGlyCysGluIleValIleAlaThrProGlyArgLeuIle 272

Qy     1282 GATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAA 1341
Db      273 AspPheLeuSerAlaGlySerThrAsnLeuLysArgCysThrTyrLeuValLeuAspGlu 292

Qy     1342 GCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCA 1401
Db      293 AlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgLysIleValSer------ 310

Qy     1402 GGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAA 1461
Db      311 ------GlnIleArgProAspArgGlnThrLeuMetTrpSerAlaThrTrpProLysGlu 328

Qy     1462 ATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAA 1521
Db      329 ValLysGlnLeuAlaGluAspPheLeu---GlyAsnTyrIleGlnIleAsnIleGlySer 347

Qy     1522 GTGGGT---GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCA 1578
Db      348 LeuGluLeuSerAlaAsnHisAsnIleArgGlnValValAspValCysAspGluPheSer 367

Qy     1579 AAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAA--------------- 1623
Db      368 LysGluGluLysLeuLysThrLeuLeuSerAspIleTyrAspThrSerGluSerProGly 387

Qy     1624 AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGT 1683
Db      388 LysIleIleIlePheValGluThrLysArgArgValAspAsnLeuValArgPheIleArg 407

Qy     1684 CAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAA 1743
Db      408 SerPheGlyValArgCysGlyAlaIleHisGlyAspLysSerGlnSerGluArgAspPhe 427

Qy     1744 GCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCC 1803
Db      428 ValLeuArgGluPheArgSerGlyLysSerAsnIleLeuValAlaThrAspValAlaAla 447

Qy     1804 AGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATT 1863
Db      448 ArgGlyLeuAspValAspGlyIleLysTyrValIleAsnPheAspTyrProGlnAsnSer 467

Qy     1864 GATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATT 1923
Db      468 GluAspTyrIleHisArgIleGlyArgThrGlyArgSerAsnThrLysGlyThrSerPhe 487

Qy     1924 TCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACA 1983
Db      488 AlaPhePhe---ThrLysAsnAsnAlaLysGlnAlaLysAlaLeuValAspValLeuArg 506

Qy     1984 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCT 2043
Db      507 GluAlaAsnGlnGluIleAsnProAlaLeuGluAsnLeuAlaArgAsnSerArgTyrAsp 526

Qy     2044 GGCTTCAGTGGTAGTACAAGA------------------GGAAACGTGTTTGCATCAGTT 2085
Db      527 GlyGlyGlyGlyProSerArgTyrGlyGlyGlyGlyGlyGlyGlyArgPheGlyGlyGly 546

Qy     2086 GATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTT 2127
Db      547 GlyPheLysLysGly---SerLeuSerAsnGlyArgGlyPhe 559

RESULT 2
US-10-228-897-13
; Sequence 13, Application US/10228897
; Publication No. US20030092043A1
; GENERAL INFORMATION:
;  APPLICANT: Fisher, Paul B.
;  APPLICANT:  Kang, Dong-Chul
;  APPLICANT:  Gopalkrishnan, Rahul V.
;  TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
;  TITLE OF INVENTION:  GENE-5 AND PROMOTER AND USES THEREOF
;  FILE REFERENCE: A34614-A-PCT-USA (070050.2121)
;  CURRENT APPLICATION NUMBER: US/10/228,897
;  CURRENT FILING DATE:  2002-08-26
;  PRIOR APPLICATION NUMBER: PCT/US01/06960
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 09/515,363
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 25
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 500
;   TYPE: PRT
;   ORGANISM: molgula oculata
US-10-228-897-13

Alignment Scores:
Pred. No.:               2.18e-57      Length:         500
Score:                   761.00        Matches:        167
Percent Similarity:      58.22%        Conservative:   88
Best Local Similarity:   38.13%        Mismatches:     125
Query Match:             19.25%        Indels:         58
DB:                      9             Gaps:           11

US-09-714-865-15 (1-2172) x US-10-228-897-13 (1-500)

Qy        826 GTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTC 885
              :::   |||   ||||||:::   |||     :::|||   |||   ||||||||
Db         82 IleThrValArgGlyHisAsnCysProLysProValLeuAsnPheTyrGluAlaAsnPhe 101

Qy        886 TGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAA 945
                       :::   :::   ||||||:::     :::|||:::   |||   :::|||
Db        102 ProAlaAsnValMetAspValIleAlaArgGlnAsnPheThrGluProThrAlaIleGln 121

Qy        946 AAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGG 1005
                          |||:::   |||:::|||   |||::::::      |||||||||||
Db        122 AlaGlnGlyTrpProValAlaLeuSerGlyLeuAspMetValGlyValAlaGlnThrGly 141

Qy       1006 TCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCAT--------- 1056
              |||||||||||   ::::::|||||||||   :::   |||:::   |||
Db        142 SerGlyLysThrLeuSerTyrLeuLeuProAlaIleValHisIleAsnHisGlnProPhe 161

Qy       1057 ------------GATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
                          ||||||                              |||   |||
Db        162 LeuGluArgGlyAspGly------------------------------ProIleCys 170

Qy       1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164
              :::::::::||||||||||||||||||      |||:::         |||   ::::::
Db        171 LeuValLeuAlaProThrArgGluLeuAlaGlnGlnValGlnGlnValAlaAlaGluTyr 190

Qy       1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA 1224
                                :::::::::   |||||||||||         |||
Db        191 CysArgAlaCysArgLeuLysSerThrCysIleTyrGlyGlyAlaProLysGlyProGln 210

Qy       1225 ATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGAT 1284
              ||||||   :::   :::|||   |||   |||||||||||||||||:::|||
Db        211 IleArgAspLeuGluArgGlyValGluIleCysIleAlaThrProGlyArgLeuIleAsp 230

Qy       1285 ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCT 1344
                 :::      |||      |||::::::   ||||||||||||||||||||||
Db        231 PheLeuGluCysGlyLysThrAsnLeuArgArgThrThrTyrLeuValLeuAspGluAla 250

Qy       1345 GATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
              ||||||||||||||||||||||||   |||:::::::::|||::::::
Db        251 AspArgMetLeuAspMetGlyPheGluProGlnIleArgLysIleValAsp--------- 267

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
                 :::      ||||||||||||||:::|||||||||:::|||:::|||:::
Db        268 ---GlnIleArgProAspArgGlnThrLeuMetTrpSerAlaThrTrpProLysGluVal 286

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGA------ 1518
              ::::::||||||   :::|||||||||   :::|||:::   :::   :::|||
Db        287 ArgGlnLeuAlaGluAspPheLeuLys---AspTyrIleHisIleAsnIleGlyAlaLeu 305

Qy       1519 --------------------------CAAGTGGGTGGAGCATGTAGAGATGTTCAGCAG 1551
                                        |||:::      |||   ||||||:::
Db        306 GluLeuSerAlaAsnHisAsnIleLeuGlnIleValAspValCysHisAspValGlu--- 324

Qy       1552 ACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611
                                        |||   |||||||||:::   ::::::
Db        325 --------------------------LysAspGluLysLeuIleArgLeuMetGluGlu 335

Qy       1612 ATAGGGGATGAA---------AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGAT 1662
              |||      |||         :::|||:::||||||||||||||||||::::::   |||
Db        336 IleMetSerGluLysGluAsnLysThrIleValPheValGluThrLysArgArgCysAsp 355

Qy       1663 TTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGG 1722
                       :::   ::::::                     ||||||||||||:::
Db        356 GluLeuThrArgLysMetArgArgAspGlyTrpProAlaMetGlyIleHisGlyAspLys 375

Qy       1723 GAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTT 1782
                 |||:::|||||||:::      |||   :::|||:::   ||||||   |||:::|||
Db        376 SerGlnGlnGluArgAspTrpValLeuAsnGluPheLysHisGlyLysAlaProIleLeu 395

Qy       1783 GTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAAT 1842
              :::||||||   ||||||:::||||||||||||:::|||:::|||:::   |||||||||
Db        396 IleAlaThrAspValAlaSerArgGlyLeuAspValGluAspValLysPheValIleAsn 415

Qy       1843 TTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGT 1902
              :::|||   |||::::::   ::::::|||:::||||||||||||||||||      |||
Db        416 TyrAspTyrProAsnSerSerGluAspTyrIleHisArgIleGlyArgThrAlaArgSer 435

Qy       1903 GGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATTTAGCACAG 1962
                 ||||||   |||   :::||||||         :::   |||::::::   |||
Db        436 ThrLysThrGlyThrAlaTyrThrPhePhe---------ThrProAsnAsnIleLysGln 452

Qy       1963 ------CCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTT---CCTGCATGGTTG 2013
                    |||:::   ||||||   :::|||   |||   :::   |||   |||
Db        453 ValSerAspLeuIleSerValLeuArgGluAlaAsnGlnAlaIleAsnProLysLeuLeu 472

Qy       2014 GAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGAGGA 2067
              :::   :::                              ||||||   :::||||||
Db        473 GlnLeuValGlu---------------AspArgGlySerGlyArgSerArgGly 485

RESULT 3
US-09-923-831-43
; Sequence 43, Application US/09923831
; Patent No. US20020115142A1
; GENERAL INFORMATION:
;  APPLICANT: Martelange, Val,rie
;  APPLICANT:  De Smet, Charles
;  APPLICANT:  Boon-Falleur, Thierry
;  TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE: L0461/7054
;  CURRENT APPLICATION NUMBER: US/09/923,831
;  CURRENT FILING DATE:  2001-08-07
;  PRIOR APPLICATION NUMBER: 09/183,706
;  PRIOR FILING DATE: 2001-10-30
;  NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 43
;   LENGTH: 648
;   TYPE: PRT
;   ORGANISM: H. sapiens
US-09-923-831-43

Alignment Scores:
Pred. No.:               7.26e-53      Length:         648
Score:                   709.50        Matches:        160
Percent Similarity:      60.00%        Conservative:   77
Best Local Similarity:   40.51%        Mismatches:     145
Query Match:             17.94%        Indels:         13
DB:                      10            Gaps:           8

US-09-714-865-15 (1-2172) x US-09-923-831-43 (1-648)

Qy        850 CCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGT---CAGACACTGAATAACAAC 906
              |||               ||||||::::::|||      |||         :::      |||
Db        237 ProAsnProThrCysThrPheAspAspAlaPheGlnCysTyrProGluValMetGluAsn 256

Qy        907 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 966
              |||   |||||||||:::   |||   ||||||:::|||         :::   ||||||:::
Db        257 IleLysLysAlaGlyPheGlnLysProThrProIleGlnSerGlnAlaTrpProIleVal 276

Qy        967 CTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||   |||   ||||||:::         ||||||||||||:::|||||||||      :::
```

```
Db        277 LeuGlnGlyIleAspLeuIleGlyValAlaGlnThrGlyThrGlyLysThrLeuCysTyr 296

Qy       1027 CTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG 1086
              |||:::|||          |||::::::          :::      |||         :::
Db        297 LeuMetProGlyPheIleHisLeuVal---------LeuGlnProSerLeuLysGlyGln 313

Qy       1087 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
                       |||    :::::::::   ||||||||||||||        |||:::
Db        314 ArgAsnArgProGlyMetLeuValLeuThrProThrArgGluLeuAlaLeuGlnValGlu 333

Qy       1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
                 |||      |||:::|||:::          :::|||:::|||   :::|||||||||
Db        334 GlyGluCysCysLysTyrSerTyr---LysGlyLeuArgSerValCysValTyrGlyGly 352

Qy       1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
                          |||   ::::::   :::|||   :::|||:::         ||||||
Db        353 GlyAsnArgAspGluGlnIleGluGluLeuLysLysGlyValAspIleIleIleAlaThr 372

Qy       1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
              |||||||||||   |||:::                :::   ||||||  |||   |||
Db        373 ProGlyArgLeuAsnAspLeuGlnMetSerAsnPheValAsnLeuLysAsnIleThrTyr 392

Qy       1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
              |||||||||||||||||||:::|||||||||||||||||    |||::::::    |||
Db        393 LeuValLeuAspGluAlaAspLysMetLeuAspMetGlyPheGluProGlnIleMetLys 412

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              ::::::                    :::      :::   |||||||||:::|||   ||||||
Db        413 IleLeu------------LeuAspValArgProAspArgGlnThrValMetThrSerAla 428

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              |||:::|||       :::   |||||||||    :::||||||        :::
Db        429 ThrTrpProHisSerValHisArgLeuAlaGlnSerTyrLeuLysGluPro---MetIle 447

Qy       1507 GTTGCTGTTGGACAAGTGGGT---GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAA 1563
              |||   ||||||   :::              |||      |||:::|||   ::::::
Db        448 ValTyrValGlyThrLeuAspLeuValAlaValSerSerValLysGlnAsnIleIleVal 467

Qy       1564 GTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT--- 1620
                    :::      |||            :::         |||:::::::::
Db        468 ThrThrGluGluGluLysTrpSerHisMetGlnThrPheLeuGlnSerMetSerSerThr 487

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
              ::::::   :::|||||||||         |||      ||||||      ::::::   |||
Db        488 AspLysValIleValPheValSerArgLysAlaValAlaAspHisLeuSerSerAspLeu 507

Qy       1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
                          ||||||      |||:::|||||||||||||||||||||:::||||||
Db        508 IleLeuGlyAsnIleSerValGluSerLeuHisGlyAspArgGluGlnArgAspArgGlu 527

Qy       1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
              :::||||||   :::|||:::   ||||||      :::|||:::||||||    :::|||
Db        528 LysAlaLeuGluAsnPheLysThrGlyLysValArgIleLeuIleAlaThrAspLeuAla 547

Qy       1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
              :::||||||||||||:::      :::|||   ||||||   |||||||||    |||
Db        548 SerArgGlyLeuAspValHisAspValThrHisValTyrAsnPheAspPheProArgAsn 567

Qy       1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
              |||:::|||||||||||||||||||||||||||||||||   |||   ||||||    :::
Db        568 IleGluGluTyrValHisArgIleGlyArgThrGlyArgAlaGlyArgThrGlyValSer 587

Qy       1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
              |||:::            :::|||      :::|||      |||:::   :::|||
Db        588 IleThrThrLeu---ThrArgAsnAspTrpArgValAlaSerGluLeuIleAsnIleLeu 606

Qy       1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCC 2025
                    |||   |||   :::|||      |||      :::|||
Db        607 GluArgAlaAsnGlnSerIleProGluGluLeuValSerMetAla 621
```

RESULT 4
US-10-147-268-2
; Sequence 2, Application US/10147268
; Patent No. US20020143154A1
; GENERAL INFORMATION:
; APPLICANT: Kaufmann, Joerg
; TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
; FILE REFERENCE: PP-01513.104/200130.456D3
; CURRENT APPLICATION NUMBER: US/10/147,268
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-147-268-2

Alignment Scores:

Pred. No.:	2.89e-52	Length:	1261
Score:	704.00	Matches:	227
Percent Similarity:	41.98%	Conservative:	113
Best Local Similarity:	28.02%	Mismatches:	264
Query Match:	17.80%	Indels:	206
DB:	12	Gaps:	26

US-09-714-865-15 (1-2172) x US-10-147-268-2 (1-1261)

```
Qy        139 GGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGA 198
              ||||||      :::      |||            |||||||||          :::   |||
Db         52 GlyProGlyThrLysArgGlyPheGlyPheGlyGlyPheAlaIle------SerAlaGly 69

Qy        199 AACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT--------------ACATCCACA 243
                 ::::::      :::      ::::::   :::                |||||||:::
Db         70 LysLysGluGluProLysLeuProGlnGlnSerHisSerAlaPheGlyAlaThrSerSer 89

Qy        244 ATGGGTGGTTTTGGA-----------------------------------GTTGGAAAG 267
                       |||||||||                                  :::|||
Db         90 SerSerGlyPheGlyLysSerAlaProProGlnLeuProSerPheTyrLysIleGlySer 109

Qy        268 AGTTTGGAAACAGAGGTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTC 327
                       |||            |||:::   |||||||||   :::
Db        110 LysArgAlaAsnPheAspGluGluAsnAlaTyrPheGluAspGluGlu----------- 125

Qy        328 TGGAGAGAGTCTAGTAAT-----------------GACTGCGAAGATAATCCAACACGG 369
                 :::|||||||||                         |||::::::|||||||||
Db        126 ---GluAspSerSerAsnValAspLeuProTyrIleProAlaGluAsnSerProThrArg 144

Qy        370 AACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGG 429
                 :::      ||||||
Db        145 GlnGlnPheHisSerLys------------------------------------------ 150

Qy        430 CCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGA 489

Db        150 ------------------------------------------------------------ 150

Qy        490 AGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGT 549
                 ::::::|||   |||   |||
Db        151 ProValAspSerAspSerAspAspAsp--------------------------------- 159

Qy        550 TCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGA------ 603
                       |||:::   :::               :::         :::|||
Db        160 ---------ProLeuGluAlaPheMetAlaGluValGluAspGlnAlaAlaArgAspMet 176

Qy        604 ---------AGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654
                                   ||||||         ||||||:::   :::::::::
Db        177 LysArgLeuGluGluLysAspLysGluArgLysAsnValLysGlyIleArgAspAspIle 196

Qy        655 ATAACAGGCTCTGGAAAGAATTCTTGG--------------------------------- 681
                          :::   ::::::
Db        197 GluGluGluAspAspGlnGluAlaTyrPheArgTyrMetAlaGluAsnProThrAlaGly 216

Qy        682 ------AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGA----------- 723
                 :::   |||   |||            |||:::   |||
Db        217 ValValGlnGluGluGluGluAspAsnLeuGluTyrAspSerAspGlyAsnProIleAla 236
```

```
Qy     724 CCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT-------------------- 762
           |||         |||  |||  ||||||  |||
Db     237 ProThrLysLysIleIleAspProLeuProProIleAspHisSerGluIleAspTyrPro 256

Qy     763 ------GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAAC-------------- 801
                 |||  :::  :::     ::::::       |||
Db     257 ProPheGluLysAsnPheTyrAsnGluHisGluGluIleThrAsnLeuThrProGlnGln 276

Qy     802 ---TTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCA 858
              |||            :::  :::  ||||||||         |||
Db     277 LeuIleAspLeuArgHisLysLeuAsnLeuArgValSerGlyAlaAlaProProArgPro 296

Qy     859 ATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCT 918
              :::|||         :::  |||  :::  |||  |||:::
Db     297 GlySerSerPheAlaHisPheGlyPheAspGluGlnLeuMetHisGlnIleArgLysSer 316

Qy     919 GGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGA 978
              ||||||:::  ||||||:::|||       :::|||:::  |||:::|||||
Db     317 GluTyrThrGlnProThrProIleGlnCysGlnGlyValProValAlaLeuSerGlyArg 336

Qy     979 GATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
           |||::::::   |||:::|||||||||||||||||||||||||||:::  |||:::
Db     337 AspMetIleGlyIleAlaLysThrGlySerGlyLysThrAlaAlaPheIleTrpProMet 356

Qy    1039 TTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA--- 1095
           |||  |||:::|||                       ||||||||:::
Db     357 LeuIleHisIleMetAspGln--------------------LysGluLeuGluProGly 369

Qy    1096 ------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149
                 |||       :::||||||  ||||||||||||||     ||||||:::
Db     370 AspGlyProIleAlaValIleValCysProThrArgGluLeuCysGlnGlnIleHisPro 389

Qy    1150 GAAGCCAGAAAATTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACC 1209
           |||   ::::::|||             :::|||:::|||   :::|||||||||
Db     390 GluCysLysArgPheGlyLysAlaTyrAsnLeuArgSerValAlaValTyrGlyGlyGly 409

Qy    1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269
              :::          :::  :::  :::|||    |||:::       ||||||
Db     410 SerMetTrpGluGlnAlaLysAlaLeuGlnGluGlyAlaGluIleValValCysThrPro 429

Qy    1270 GGAAGACTGATGGATATCATAGGCAAAGAAAAAGATTGGTCTCAAACAGATCAAATACTTA 1329
           |||||||||:::|||   :::  |||:::       |||:::::::::  ||||||
Db     430 GlyArgLeuIleAspHisValLysLysLysAlaThrAsnLeuGlnArgValSerTyrLeu 449

Qy    1330 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTA 1389
           |||   ||||||||||||||||||  |||||||||||||     :::::::::   :::
Db     450 ValPheAspGluAlaAspArgMetPheAspMetGlyPheGluTyrGlnValArgSerIle 469

Qy    1390 ATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACT 1449
              |||             :::       ||||||||||||:::||||||||||||
Db     470 AlaSer-----------HisValArgProAspArgGlnThrLeuLeuPheSerAlaThr 485

Qy    1450 TTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTT 1509
           |||    :::::|||::::::||||||         :::   :::    |||
Db     486 PheArgLysLysIleGluLysLeuAlaArgAspIleLeu---IleAspProIleArgVal 504

Qy    1510 GCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT------CTCCAA 1563
              |||   :::|||  |||      ||||||   |||   |||      |||
Db     505 ValGlnGlyAspIleGlyGluAlaAsnGluAspValThrGlnIleValGluIleLeuHis 524

Qy    1564 GTTGGCCAGTTCTCAAAA-----------AGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611
              |||      ||||||                  :::|||||||||
Db     525 SerGlyPro---SerLysTrpAsnTrpLeuThrArgArgLeuValGluPhe--------- 540

Qy    1612 ATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCA 1671
                       ::::::||||||       |||       |||:::       |||
Db     541 ThrSerSerGlySerValLeuLeuPheValThrLysLysAlaAsnAlaGluGluLeuAla 560

Qy    1672 ACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGA 1731
              |||   ||||||          :::       :::|||||||||   :::|||
Db     561 AsnAsnLeuLysGlnGluGlyHisAsnLeuGlyLeuLeuHisGlyAspMetAspGlnSer 580

Qy    1732 GAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACT 1791
           ||||||   :::   :::   ||||||:::            ||||||||||||||||||
Db     581 GluArgAsnLysValIleSerAspPheLysLysLysAspIleProValLeuValAlaThr 600

Qy    1792 TCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTT 1851
              ||||||||||||||||||||||||   :::::::::   |||||||||:::|||:::
Db     601 AspValAlaAlaArgGlyLeuAspIleProSerIleLysThrValIleAsnTyrAspVal 620

Qy    1852 CCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACT 1911
                 ||||||   :::   ||||||||||||||||||||||||||   |||
Db     621 AlaArgAspIleAspThrHisThrHisArgIleGlyArgThrGlyArgAlaGlyGluLys 640

Qy    1912 GGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTA 1971
           |||   |||   :::                  |||::::::   |||    ||||||
Db     641 GlyValAlaTyrThrLeuLeu---ThrProLysAspSerAsnPheAlaGlyAspLeuVal 659

Qy    1972 AAAGTATTGACAGATGCTCAACAGGATGTTCCT-------------------------- 2004
              :::   |||         |||   |||   |||
Db     660 ArgAsnLeuGluGlyAlaAsnGlnHisValSerLysGluLeuLeuAspLeuAlaMetGln 679

Qy    2005 ---GCATGGTTGGAAGAAATTGCCTTTAGT---------------------------- 2031
              ||||||      :::      |||
Db     680 AsnAlaTrpPheArgLysSerArgPheLysGlyGlyLysGlyLysLysLeuAsnIleGly 699

Qy    2032 -----------------ACATACATTCCTGGCTTCAGTGGTAGTACA------AGAGGA 2067
                                ||||||                           ||||||
Db     700 GlyGlyGlyLeuGlyTyrArgGluArgProGlyLeuGlySerGluAsnMetAspArgGly 719

Qy    2068 AAC--------------------------------------------------------- 2070
           |||
Db     720 AsnAsnAsnValMetSerAsnTyrGluAlaTyrLysProSerThrGlyAlaMetGlyAsp 739

Qy    2071 ---------------------------------------------------GTGTTTGCA 2079
                                                              |||   |||
Db     740 ArgLeuThrAlaMetLysAlaAlaPheGlnSerGlnTyrLysSerHisPheValAlaAla 759

Qy    2080 TCAGTTGATACCAGAAAG---GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCA 2136
           |||:::         :::|||   |||   |||               :::|||::::::|||:::
Db     760 SerLeuSerAsnGlnLysAlaGlySerSerAlaAlaGlyAlaSerGlyTrpThrSerAla 779

Qy    2137 CGAGCTCCCAATCCAGTAGATGATGAGTCA 2166
              :::   |||   |||         |||
Db     780 GlySerLeuAsnSerValProThrAsnSer 789
```

```
RESULT 5
US-09-736-457-335
; Sequence 335, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 335
;   LENGTH: 709
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-736-457-335

Alignment Scores:
Pred. No.:              2.05e-50       Length:        709
```

```
Score:                    681.50     Matches:        195
Percent Similarity:       45.98%     Conservative:   91
Best Local Similarity:    31.35%     Mismatches:     213
Query Match:              17.24%     Indels:         123
DB:                       9          Gaps:           18

US-09-714-865-15 (1-2172) x US-09-736-457-335 (1-709)

Qy        619 GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT 678
              ||||||         ||||||:::  :::::::::             :::   :::
Db         21 GluArgLysAsnValLysGlyIleArgAspAspIleGluGluGluAspAspGlnGluAla 40

Qy        679 TGG---------------------------------------AAGTCAGAAGCAGAAGGA 699
              :::                                       :::   |||   |||
Db         41 TyrPheArgTyrMetAlaGluAsnProThrAlaGlyValValGlnGluGluGluGluAsp 60

Qy        700 GGAGAAAGTAGTGATACTCAAGGA------------CCAAAAGTGACCTACATACCCCCT 747
                       |||:::   |||               |||         |||   |||
Db         61 AsnLeuGluTyrAspSerAspGlyAsnProIleAlaProThrLysLysIleIleAspPro 80

Qy        748 CCTCCACCTGAGGAT---------------------------GAGGACTCCATCTTTGCA 780
                 ||||||   |||                           |||   :::   :::
Db         81 LeuProProIleAspHisSerGluIleAspTyrProProPheGluLysAsnPheTyrAsn 100

Qy        781 CATTATCAGACAGGCATAAAC------------------TTCGACAAATACGACACTATT 822
                 ::::::      |||                     |||               :::
Db        101 GluHisGluGluIleThrAsnLeuThrProGlnGlnLeuIleAspLeuArgHisLysLeu 120

Qy        823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                 :::   |||||||||         |||         :::|||
Db        121 AsnLeuArgValSerGlyAlaAlaProProArgProGlySerSerPheAlaHisPheGly 140

Qy        883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
                 :::   |||   :::   |||   |||:::   ||||||:::   ||||||:::
Db        141 PheAspGluGlnLeuMetHisGlnIleArgLysSerGluTyrThrGlnProThrProIle 160

Qy        943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
              |||         :::|||:::   |||:::|||||||||:::::     |||:::|||
Db        161 GlnCysGlnGlyValProValAlaLeuSerGlyArgAspMetIleGlyIleAlaLysThr 180

Qy       1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA 1062
              ||||||||||||||||||||||||:::   |||:::|||   |||:::|||
Db        181 GlySerGlyLysThrAlaAlaPheIleTrpProMetLeuIleHisIleMetAspGln--- 199

Qy       1063 ATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA---------CCAGAGTGTATTATTGTA 1113
                                   |||||||||:::         |||      :::||||||
Db        200 ------------------LysGluLeuGluProGlyAspGlyProIleAlaValIleVal 213

Qy       1114 GCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGG 1173
                 |||||||||||||||         ||||||:::   |||   ::::::|||
Db        214 CysProThrArgGluLeuCysGlnGlnIleHisAlaGluCysLysArgPheGlyLysAla 233

Qy       1174 ACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAA 1233
                 :::|||:::|||   :::|||||||||         :::         :::
Db        234 TyrAsnLeuArgSerValAlaValTyrGlyGlyGlySerMetTrpGluGlnAlaLysAla 253

Qy       1234 ATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGC 1293
                 :::   :::|||      |||:::      |||||||||||||||:::|||
Db        254 LeuGlnGluGlyAlaGluIleValValCysThrProGlyArgLeuIleAspHisValLys 273

Qy       1294 AAAGAAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATG 1353
              ||||:::      |||:::::::::   ||||||||   ||||||||||||||||||
Db        274 LysLysAlaThrAsnLeuGlnArgValSerTyrLeuValPheAspGluAlaAspArgMet 293

Qy       1354 TTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCA 1413
                 ||||||||||||      :::::::::   :::   |||
Db        294 PheAspMetGlyPheGluTyrGlnValArgSerIleAlaSer------------HisVal 309

Qy       1414 AAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTG 1473
              :::   |||||||||||:::|||||||||||||||      ::::::|||::::::|||
Db        310 ArgProAspArgGlnThrLeuLeuPheSerAlaThrPheArgLysLysIleGluLysLeu 329

Qy       1474 GCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCA 1533
              |||   :::   |||         :::   :::   |||         |||   :::|||   |||
Db        330 AlaArgAspIleLeu---IleAspProIleArgValValGlnGlyAspIleGlyGluAla 348

Qy       1534 TGTAGAGATGTTCAGCAGACCGTT------CTCCAAGTTGGCCAGTTCTCAAAA------ 1581
                 ||||||   |||   |||         |||         |||         ||||||
Db        349 AsnGluAspValThrGlnIleValGluIleLeuHisSerGlyPro---SerLysTrpAsn 367

Qy       1582 ------AGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTC 1635
                 :::|||||||||                                    ::::::
Db        368 TrpLeuThrArgArgLeuValGluPhe---------ThrSerSerGlySerValLeuLeu 384

Qy       1636 TTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATA 1695
              ||||||   |||   |||:::      |||         |||   ||||||
Db        385 PheValThrLysLysAlaAsnAlaGluGluLeuAlaAsnAsnLeuLysGlnGluGlyHis 404

Qy       1696 TCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGAT 1755
              :::         :::||||||||||   :::|||   ||||||   :::   :::   |||
Db        405 AsnLeuGlyLeuLeuHisGlyAspMetAspGlnSerGluArgAsnLysValIleSerAsp 424

Qy       1756 TTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGAT 1815
              |||:::            ||||||||||||||||||   ||||||||||||||||||||||
Db        425 PheLysLysLysAspIleProValLeuValAlaThrAspValAlaAlaArgGlyLeuAsp 444

Qy       1816 ATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTT 1875
              |||   :::::::::   |||||||||:::|||:::         ||||||   :::
Db        445 IleProSerIleLysThrValIleAsnTyrAspValAlaArgAspIleAspThrHisThr 464

Qy       1876 CATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGAT 1935
              ||||||||||||||||||||||||   |||         |||   |||   :::
Db        465 HisArgIleGlyArgThrGlyArgAlaGlyGluLysGlyValAlaTyrThrLeuLeu--- 483

Qy       1936 CTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAG 1995
                          |||::::::   |||      ||||||:::   |||      |||   |||
Db        484 ThrProLysAspSerAsnPheAlaGlyAspLeuValArgAsnLeuGluGlyAlaAsnGln 503

Qy       1996 GATGTTCCT---------------------------------GCATGGTTGGAAGAAATTGCC 2025
                 |||                                 ||||||      :::
Db        504 HisValSerLysGluLeuLeuAspLeuAlaMetGlnAsnAlaTrpPheArgLysSerArg 523

Qy       2026 TTTAGT------------------------------------------------ACATAC 2037
              |||
Db        524 PheLysGlyGlyLysGlyLysLysLeuAsnIleGlyGlyGlyGlyLeuGlyTyrArgGlu 543

Qy       2038 ATTCCTGGCTTCAGTGGTAGTACA------AGAGGAAAC--------------------- 2070
                 ||||||               |||||||||
Db        544 ArgProGlyLeuGlySerGluAsnMetAspArgGlyAsnAsnAsnValMetSerAsnTyr 563

Qy       2070 ------------------------------------------------------------ 2070

Db        564 GluAlaTyrLysProSerThrGlyAlaMetGlyAspArgLeuThrAlaMetLysAlaAla 583

Qy       2071 ---------------------------GTGTTTGCATCAGTTGATACCAGAAAG---GGC 2100
                                         |||   ||||||:::      :::|||   |||
Db        584 PheGlnSerGlnTyrLysSerHisPheValAlaAlaSerLeuSerAsnGlnLysAlaGly 603

Qy       2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
                 |||            :::|||::::::|||:::   :::   |||   |||
Db        604 SerSerAlaAlaGlyAlaSerGlyTrpThrSerAlaGlySerLeuAsnSerValProThr 623

Qy       2161 GAGTCA 2166
                 |||
Db        624 AsnSer 625

RESULT 6
US-09-902-941-335
; Sequence 335, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
```

```
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C17
;  CURRENT APPLICATION NUMBER: US/09/902,941
;  CURRENT FILING DATE:  2001-07-10
;  NUMBER OF SEQ ID NOS: 2002
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 335
;   LENGTH: 709
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-902-941-335

Alignment Scores:
Pred. No.:              2.05e-50        Length:         709
Score:                  681.50          Matches:        195
Percent Similarity:     45.98%          Conservative:   91
Best Local Similarity:  31.35%          Mismatches:     213
Query Match:            17.24%          Indels:         123
DB:                     9               Gaps:           18

US-09-714-865-15 (1-2172) x US-09-902-941-335 (1-709)

Qy        619 GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT 678
              ||||||          ||||||:::  :::::::::               :::   :::
Db         21 GluArgLysAsnValLysGlyIleArgAspAspIleGluGluGluAspAspGlnGluAla 40

Qy        679 TGG---------------------------------------AAGTCAGAAGCAGAAGGA 699
              :::                                       :::   |||   |||
Db         41 TyrPheArgTyrMetAlaGluAsnProThrAlaGlyValValGlnGluGluGluGluAsp 60

Qy        700 GGAGAAAGTAGTGATACTCAAGGA------------CCAAAAGTGACCTACATACCCCCT 747
                       |||:::   |||            |||            |||   |||
Db         61 AsnLeuGluTyrAspSerAspGlyAsnProIleAlaProThrLysLysIleIleAspPro 80

Qy        748 CCTCCACCTGAGGAT---------------------------GAGGACTCCATCTTTGCA 780
                 ||||||   |||                           |||   :::   :::
Db         81 LeuProProIleAspHisSerGluIleAspTyrProProPheGluLysAsnPheTyrAsn 100

Qy        781 CATTATCAGACAGGCATAAAC------------------TTCGACAAATACGACACTATT 822
                 ::::::      |||                     |||               :::
Db        101 GluHisGluGluIleThrAsnLeuThrProGlnGlnLeuIleAspLeuArgHisLysLeu 120

Qy        823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                 :::   |||||||||               |||            :::|||
Db        121 AsnLeuArgValSerGlyAlaAlaProProArgProGlySerSerPheAlaHisPheGly 140

Qy        883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
                    :::   |||   :::   |||   |||:::   ||||||:::   ||||||:::
Db        141 PheAspGluGlnLeuMetHisGlnIleArgLysSerGluTyrThrGlnProThrProIle 160

Qy        943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
              |||            :::|||:::   |||:::|||||||||::::::      |||:::|||
Db        161 GlnCysGlnGlyValProValAlaLeuSerGlyArgAspMetIleGlyIleAlaLysThr 180

Qy       1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA 1062
              ||||||||||||||||||||||||:::   |||:::|||   |||:::|||
Db        181 GlySerGlyLysThrAlaAlaPheIleTrpProMetLeuIleHisIleMetAspGln--- 199

Qy       1063 ATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA---------CCAGAGTGTATTATTGTA 1113
                                 |||||||||:::            |||      :::||||||
Db        200 ------------------LysGluLeuGluProGlyAspGlyProIleAlaValIleVal 213

Qy       1114 GCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGG 1173
                 ||||||||||||||      ||||||:::   |||   ::::::|||
Db        214 CysProThrArgGluLeuCysGlnGlnIleHisAlaGluCysLysArgPheGlyLysAla 233

Qy       1174 ACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGGGACATTCAATTCGACAA 1233
                 :::|||:::|||   :::|||||||||         :::               :::
Db        234 TyrAsnLeuArgSerValAlaValTyrGlyGlyGlySerMetTrpGluGlnAlaLysAla 253

Qy       1234 ATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGC 1293
              :::   :::|||         |||:::   ||||||||||||||:::|||   :::
Db        254 LeuGlnGluGlyAlaGluIleValValCysThrProGlyArgLeuIleAspHisValLys 273

Qy       1294 AAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATG 1353
              |||:::         |||:::::::::   |||||||||   ||||||||||||||||||
Db        274 LysLysAlaThrAsnLeuGlnArgValSerTyrLeuValPheAspGluAlaAspArgMet 293

Qy       1354 TTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCA 1413
                 ||||||||||||      :::::::::   :::   |||
Db        294 PheAspMetGlyPheGluTyrGlnValArgSerIleAlaSer-----------HisVal 309

Qy       1414 AAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTG 1473
              :::      ||||||||||||:::||||||||||||||||||   ::::::|||::::::|||
Db        310 ArgProAspArgGlnThrLeuLeuPheSerAlaThrPheArgLysLysIleGluLysLeu 329

Qy       1474 GCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCA 1533
              |||   :::   |||      :::   :::   |||         |||   :::|||   |||
Db        330 AlaArgAspIleLeu---IleAspProIleArgValValGlnGlyAspIleGlyGluAla 348

Qy       1534 TGTAGAGATGTTCAGCAGACCGTT------CTCCAAGTTGGCCAGTTCTCAAAA------ 1581
                 ||||||   |||   |||         |||      |||      ||||||
Db        349 AsnGluAspValThrGlnIleValGluIleLeuHisSerGlyPro---SerLysTrpAsn 367

Qy       1582 ------AGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTC 1635
                    :::|||||||||                              ::::::
Db        368 TrpLeuThrArgArgLeuValGluPhe---------ThrSerSerGlySerValLeuLeu 384

Qy       1636 TTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATA 1695
              ||||||   |||      |||:::      |||      |||      ||||||
Db        385 PheValThrLysLysAlaAsnAlaGluGluLeuAlaAsnAsnLeuLysGlnGluGlyHis 404

Qy       1696 TCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGAT 1755
              :::      :::|||||||||   :::|||   ||||||   :::   :::   |||
Db        405 AsnLeuGlyLeuLeuHisGlyAspMetAspGlnSerGluArgAsnLysValIleSerAsp 424

Qy       1756 TTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGAT 1815
              |||:::      ||||||||||||||||||   ||||||||||||||||||||||||
Db        425 PheLysLysLysAspIleProValLeuValAlaThrAspValAlaAlaArgGlyLeuAsp 444

Qy       1816 ATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTT 1875
              |||   :::::::::   |||||||||:::|||:::               ||||||   :::
Db        445 IleProSerIleLysThrValIleAsnTyrAspValAlaArgAspIleAspThrHisThr 464

Qy       1876 CATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGAT 1935
              ||||||||||||||||||||||||   |||      |||   |||   :::
Db        465 HisArgIleGlyArgThrGlyArgAlaGlyGluLysGlyValAlaTyrThrLeuLeu--- 483

Qy       1936 CTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAG 1995
                 |||::::::   |||      ||||||:::   |||         |||   |||
Db        484 ThrProLysAspSerAsnPheAlaGlyAspLeuValArgAsnLeuGluGlyAlaAsnGln 503

Qy       1996 GATGTTCCT-----------------------------GCATGGTTGGAAGAAATTGCC 2025
                 |||                                  ||||||   :::
Db        504 HisValSerLysGluLeuLeuAspLeuAlaMetGlnAsnAlaTrpPheArgLysSerArg 523

Qy       2026 TTTAGT------------------------------------------------ACATAC 2037
              |||
Db        524 PheLysGlyGlyLysGlyLysLysLeuAsnIleGlyGlyGlyGlyLeuGlyTyrArgGlu 543

Qy       2038 ATTCCTGGCTTCAGTGGTAGTACA------AGAGGAAAC--------------------- 2070
                 ||||||                     |||||||||
Db        544 ArgProGlyLeuGlySerGluAsnMetAspArgGlyAsnAsnAsnValMetSerAsnTyr 563

Qy       2070 ------------------------------------------------------------ 2070

Db        564 GluAlaTyrLysProSerThrGlyAlaMetGlyAspArgLeuThrAlaMetLysAlaAla 583

Qy       2071 --------------------------GTGTTTGCATCAGTTGATACCAGAAAG---GGC 2100
```

```
                                         |||   ||||||:::      :::|||   |||
Db        584 PheGlnSerGlnTyrLysSerHisPheValAlaAlaSerLeuSerAsnGlnLysAlaGly 603

Qy       2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
                 |||         :::|||::::::|||:::   :::   |||   |||
Db        604 SerSerAlaAlaGlyAlaSerGlyTrpThrSerAlaGlySerLeuAsnSerValProThr 623

Qy       2161 GAGTCA 2166
                 |||
Db        624 AsnSer 625

RESULT 7
US-09-849-626-335
; Sequence 335, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Switzer, Anne
;  APPLICANT:  McNeill, Patricia
;  APPLICANT:  Clapper, Jonathan
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C16
;  CURRENT APPLICATION NUMBER: US/09/849,626
;  CURRENT FILING DATE:  2001-05-03
;  NUMBER OF SEQ ID NOS: 1926
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 335
;   LENGTH: 709
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-849-626-335

Alignment Scores:
Pred. No.:                2.05e-50     Length:        709
Score:                    681.50       Matches:       195
Percent Similarity:       45.98%       Conservative:  91
Best Local Similarity:    31.35%       Mismatches:    213
Query Match:              17.24%       Indels:        123
DB:                       9            Gaps:          18

US-09-714-865-15 (1-2172) x US-09-849-626-335 (1-709)

Qy        619 GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT 678
                ||||||         ||||||:::  :::::::::          :::   :::
Db         21 GluArgLysAsnValLysGlyIleArgAspAspIleGluGluGluAspAspGlnGluAla 40

Qy        679 TGG---------------------------------------AAGTCAGAAGCAGAAGGA 699
              :::                                       :::   |||   |||
Db         41 TyrPheArgTyrMetAlaGluAsnProThrAlaGlyValValGlnGluGluGluGluAsp 60

Qy        700 GGAGAAAGTAGTGATACTCAAGGA------------CCAAAAGTGACCTACATACCCCCT 747
                          |||:::   |||            |||         |||   |||
Db         61 AsnLeuGluTyrAspSerAspGlyAsnProIleAlaProThrLysLysIleIleAspPro 80

Qy        748 CCTCCACCTGAGGAT---------------------------GAGGACTCCATCTTTGCA 780
                 ||||||   |||                           |||   :::   :::
Db         81 LeuProProIleAspHisSerGluIleAspTyrProProPheGluLysAsnPheTyrAsn 100

Qy        781 CATTATCAGACAGGCATAAAC------------------TTCGACAAATACGACACTATT 822
                 ::::::         |||                  |||               :::
Db        101 GluHisGluGluIleThrAsnLeuThrProGlnGlnLeuIleAspLeuArgHisLysLeu 120

Qy        823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                 :::   |||||||||         |||            :::|||
Db        121 AsnLeuArgValSerGlyAlaAlaProProArgProGlySerSerPheAlaHisPheGly 140

Qy        883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
                 :::   |||   :::   |||   |||:::   ||||||:::   ||||||:::
Db        141 PheAspGluGlnLeuMetHisGlnIleArgLysSerGluTyrThrGlnProThrProIle 160

Qy        943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
              |||        :::|||:::   |||:::||||||||||||::::::      |||:::|||
Db        161 GlnCysGlnGlyValProValAlaLeuSerGlyArgAspMetIleGlyIleAlaLysThr 180

Qy       1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTGGCTCATATGATGCATGATGGA 1062
              |||||||||||||||||||||||:::   |||:::|||   |||:::|||
Db        181 GlySerGlyLysThrAlaAlaPheIleTrpProMetLeuIleHisIleMetAspGln--- 199

Qy       1063 ATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA---------CCAGAGTGTATTATTGTA 1113
                                 |||||||||:::         |||      :::||||||
Db        200 -----------------LysGluLeuGluProGlyAspGlyProIleAlaValIleVal 213

Qy       1114 GCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGG 1173
                 ||||||||||||||      ||||||:::   |||   ::::::|||
Db        214 CysProThrArgGluLeuCysGlnGlnIleHisAlaGluCysLysArgPheGlyLysAla 233

Qy       1174 ACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGGGACATTCAATTCGACAA 1233
                 :::|||:::|||   :::|||||||||         :::               :::
Db        234 TyrAsnLeuArgSerValAlaValTyrGlyGlyGlySerMetTrpGluGlnAlaLysAla 253

Qy       1234 ATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGC 1293
              :::   :::|||      |||:::      ||||||||||||||:::|||   :::
Db        254 LeuGlnGluGlyAlaGluIleValValCysThrProGlyArgLeuIleAspHisValLys 273

Qy       1294 AAAGAAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATG 1353
              |||:::            |||:::::::::   |||||||||   ||||||||||||||||
Db        274 LysLysAlaThrAsnLeuGlnArgValSerTyrLeuValPheAspGluAlaAspArgMet 293

Qy       1354 TTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCA 1413
                 ||||||||||||         :::::::::   :::   |||
Db        294 PheAspMetGlyPheGluTyrGlnValArgSerIleAlaSer-----------HisVal 309

Qy       1414 AAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTG 1473
              :::      ||||||||||||:::||||||||||||||||   ::::::|||::::::|||
Db        310 ArgProAspArgGlnThrLeuLeuPheSerAlaThrPheArgLysLysIleGluLysLeu 329

Qy       1474 GCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCA 1533
              |||   :::   |||         :::   :::   |||         |||   :::|||   |||
Db        330 AlaArgAspIleLeu---IleAspProIleArgValValGlnGlyAspIleGlyGluAla 348

Qy       1534 TGTAGAGATGTTCAGCAGACCGTT------CTCCAAGTTGGCCAGTTCTCAAAA------ 1581
                 ||||||   |||   |||         |||         |||      ||||||
Db        349 AsnGluAspValThrGlnIleValGluIleLeuHisSerGlyPro---SerLysTrpAsn 367

Qy       1582 ------AGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTC 1635
                    :::|||||||||                                  ::::::
Db        368 TrpLeuThrArgArgLeuValGluPhe---------ThrSerSerGlySerValLeuLeu 384

Qy       1636 TTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATA 1695
              ||||||      |||      |||:::         |||         |||   ||||||
Db        385 PheValThrLysLysAlaAsnAlaGluGluLeuAlaAsnAsnLeuLysGlnGluGlyHis 404

Qy       1696 TCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGAT 1755
              :::         :::|||||||||   :::|||   ||||||   :::   :::   |||
Db        405 AsnLeuGlyLeuLeuHisGlyAspMetAspGlnSerGluArgAsnLysValIleSerAsp 424

Qy       1756 TTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGAT 1815
              |||:::            ||||||||||||||||||   |||||||||||||||||||||
Db        425 PheLysLysLysAspIleProValLeuValAlaThrAspValAlaAlaArgGlyLeuAsp 444

Qy       1816 ATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTT 1875
              |||   :::::::::   |||||||||:::|||:::               ||||||   :::
Db        445 IleProSerIleLysThrValIleAsnTyrAspValAlaArgAspIleAspThrHisThr 464

Qy       1876 CATCGAATTGGGCGTACTGGTCGTTGTGGAATACTGGCAGAGCAATTTCCTTTTTTGAT 1935
              ||||||||||||||||||||||||   |||      |||   |||   :::
Db        465 HisArgIleGlyArgThrGlyArgAlaGlyGluLysGlyValAlaTyrThrLeuLeu--- 483

Qy       1936 CTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAG 1995
                 |||::::::   |||      ||||||:::   |||         |||   |||
Db        484 ThrProLysAspSerAsnPheAlaGlyAspLeuValArgAsnLeuGluGlyAlaAsnGln 503
```

```
Qy       1996 GATGTTCCT-----------------------------GCATGGTTGGAAGAAATTGCC 2025
                |||                                  |||||     :::
Db        504 HisValSerLysGluLeuLeuAspLeuAlaMetGlnAsnAlaTrpPheArgLysSerArg 523

Qy       2026 TTTAGT--------------------------------------------------ACATAC 2037
              |||
Db        524 PheLysGlyGlyLysGlyLysLysLeuAsnIleGlyGlyGlyGlyLeuGlyTyrArgGlu 543

Qy       2038 ATTCCTGGCTTCAGTGGTAGTACA------AGAGGAAAC-------------------- 2070
                 ||||||                        ||||||||
Db        544 ArgProGlyLeuGlySerGluAsnMetAspArgGlyAsnAsnAsnValMetSerAsnTyr 563

Qy       2070 ------------------------------------------------------------ 2070

Db        564 GluAlaTyrLysProSerThrGlyAlaMetGlyAspArgLeuThrAlaMetLysAlaAla 583

Qy       2071 ---------------------------GTGTTTGCATCAGTTGATACCAGAAAG---GGC 2100
                                         |||  ||||||:::   :::|||   |||
Db        584 PheGlnSerGlnTyrLysSerHisPheValAlaAlaSerLeuSerAsnGlnLysAlaGly 603

Qy       2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
                 |||             :::|||::::::|||:::   :::   |||   |||
Db        604 SerSerAlaAlaGlyAlaSerGlyTrpThrSerAlaGlySerLeuAsnSerValProThr 623

Qy       2161 GAGTCA 2166
                 |||
Db        624 AsnSer 625
```

RESULT 8
US-10-017-754-335
; Sequence 335, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 335
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-754-335

```
Alignment Scores:
Pred. No.:               2.05e-50       Length:          709
Score:                   681.50         Matches:         195
Percent Similarity:      45.98%         Conservative:    91
Best Local Similarity:   31.35%         Mismatches:      213
Query Match:             17.24%         Indels:          123
DB:                      9              Gaps:            18
```

US-09-714-865-15 (1-2172) x US-10-017-754-335 (1-709)

```
Qy        619 GAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCT 678
                 ||||||         ||||||:::   :::::::::            :::   :::
Db         21 GluArgLysAsnValLysGlyIleArgAspAspIleGluGluGluAspAspGlnGluAla 40

Qy        679 TGG---------------------------------------AAGTCAGAAGCAGAAGGA 699
              :::                                       :::   |||   |||
Db         41 TyrPheArgTyrMetAlaGluAsnProThrAlaGlyValValGlnGluGluGluAsp 60

Qy        700 GGAGAAAGTAGTGATACTCAAGGA------------CCAAAAGTGACCTACATACCCCCT 747
                          |||:::   |||             |||          |||   |||
Db         61 AsnLeuGluTyrAspSerAspGlyAsnProIleAlaProThrLysLysIleIleAspPro 80

Qy        748 CCTCCACCTGAGGAT--------------------------GAGGACTCCATCTTTGCA 780
                 ||||||   |||                          |||   :::   :::
Db         81 LeuProProIleAspHisSerGluIleAspTyrProProPheGluLysAsnPheTyrAsn 100

Qy        781 CATTATCAGACAGGCATAAAC------------------TTCGACAAATACGACACTATT 822
                 ::::::   |||                          |||          :::
Db        101 GluHisGluGluIleThrAsnLeuThrProGlnGlnLeuIleAspLeuArgHisLysLeu 120

Qy        823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                 :::   |||||||||         |||          :::|||
Db        121 AsnLeuArgValSerGlyAlaAlaProProArgProGlySerSerPheAlaHisPheGly 140

Qy        883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
                 :::   |||   :::   |||   |||:::   ||||||:::   ||||||:::
Db        141 PheAspGluGlnLeuMetHisGlnIleArgLysSerGluTyrThrGlnProThrProIle 160

Qy        943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
              |||         :::|||:::   |||:::||||||||||::::::      |||:::|||
Db        161 GlnCysGlnGlyValProValAlaLeuSerGlyArgAspMetIleGlyIleAlaLysThr 180

Qy       1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA 1062
              ||||||||||||||||||||||||:::   |||:::|||   |||:::|||
Db        181 GlySerGlyLysThrAlaAlaPheIleTrpProMetLeuIleHisIleMetAspGln--- 199

Qy       1063 ATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA---------CCAGAGTGTATTATTGTA 1113
                                 |||||||||:::          |||      :::||||||
Db        200 ------------------LysGluLeuGluProGlyAspGlyProIleAlaValIleVal 213

Qy       1114 GCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGG 1173
                 ||||||||||||||         ||||||:::   |||   ::::::|||
Db        214 CysProThrArgGluLeuCysGlnGlnIleHisAlaGluCysLysArgPheGlyLysAla 233

Qy       1174 ACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGGGACATTCAATTCGACAA 1233
                 :::|||:::|||   :::||||||||         :::         :::
Db        234 TyrAsnLeuArgSerValAlaValTyrGlyGlyGlySerMetTrpGluGlnAlaLysAla 253

Qy       1234 ATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGC 1293
              :::   :::|||      |||:::      ||||||||||||||:::|||   :::
Db        254 LeuGlnGluGlyAlaGluIleValValCysThrProGlyArgLeuIleAspHisValLys 273

Qy       1294 AAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATG 1353
              |||:::         |||:::::::::   ||||||   |||||||||||||||
Db        274 LysLysAlaThrAsnLeuGlnArgValSerTyrLeuValPheAspGluAlaAspArgMet 293

Qy       1354 TTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCA 1413
                 |||||||||||      :::::::::   :::   |||
Db        294 PheAspMetGlyPheGluTyrGlnValArgSerIleAlaSer------------HisVal 309

Qy       1414 AAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTG 1473
              :::      ||||||||||:::||||||||||||      ::::::|||::::::|||
Db        310 ArgProAspArgGlnThrLeuLeuPheSerAlaThrPheArgLysLysIleGluLysLeu 329

Qy       1474 GCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCA 1533
              |||   :::   |||         :::   :::   |||      |||   :::|||   |||
Db        330 AlaArgAspIleLeu---IleAspProIleArgValValGlnGlyAspIleGlyGluAla 348

Qy       1534 TGTAGAGATGTTCAGCAGACCGTT------CTCCAAGTTGGCCAGTTCTCAAAA------ 1581
                 ||||||   |||   |||         |||         ||||||
Db        349 AsnGluAspValThrGlnIleValGluIleLeuHisSerGlyPro---SerLysTrpAsn 367

Qy       1582 ------AGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAGAACTATGGTC 1635
                    :::|||||||||                               ::::::
Db        368 TrpLeuThrArgArgLeuValGluPhe---------ThrSerSerGlySerValLeuLeu 384

Qy       1636 TTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATA 1695
              ||||||      |||      |||:::      |||      |||   ||||||
Db        385 PheValThrLysLysAlaAsnAlaGluGluLeuAlaAsnAsnLeuLysGlnGluGlyHis 404
```

```
Qy       1696 TCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGAT 1755
              :::          :::|||||||||    :::|||   ||||||   :::    :::   |||
Db        405 AsnLeuGlyLeuLeuHisGlyAspMetAspGlnSerGluArgAsnLysValIleSerAsp 424

Qy       1756 TTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGAT 1815
              |||:::                  |||||||||||||||   ||||||||||||||||||
Db        425 PheLysLysLysAspIleProValLeuValAlaThrAspValAlaAlaArgGlyLeuAsp 444

Qy       1816 ATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTT 1875
              |||   :::::::::   ||||||||:::|||:::             ||||||   :::
Db        445 IleProSerIleLysThrValIleAsnTyrAspValAlaArgAspIleAspThrHisThr 464

Qy       1876 CATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGAT 1935
              ||||||||||||||||||||||||   |||      |||   |||   :::
Db        465 HisArgIleGlyArgThrGlyArgAlaGlyGluLysGlyValAlaTyrThrLeuLeu--- 483

Qy       1936 CTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAG 1995
                    |||::::::   |||      ||||||:::   |||         |||   |||
Db        484 ThrProLysAspSerAsnPheAlaGlyAspLeuValArgAsnLeuGluGlyAlaAsnGln 503

Qy       1996 GATGTTCCT------------------------------GCATGGTTGGAAGAAATTGCC 2025
                 |||                                 ||||||   :::
Db        504 HisValSerLysGluLeuLeuAspLeuAlaMetGlnAsnAlaTrpPheArgLysSerArg 523

Qy       2026 TTTAGT---------------------------------------------------ACATAC 2037
              |||
Db        524 PheLysGlyGlyLysGlyLysLysLeuAsnIleGlyGlyGlyGlyLeuGlyTyrArgGlu 543

Qy       2038 ATTCCTGGCTTCAGTGGTAGTACA------AGAGGAAAC--------------------- 2070
                 ||||||                     |||||||||
Db        544 ArgProGlyLeuGlySerGluAsnMetAspArgGlyAsnAsnAsnValMetSerAsnTyr 563

Qy       2070 ------------------------------------------------------------ 2070

Db        564 GluAlaTyrLysProSerThrGlyAlaMetGlyAspArgLeuThrAlaMetLysAlaAla 583

Qy       2071 ---------------------------GTGTTTGCATCAGTTGATACCAGAAAG---GGC 2100
                                         |||   ||||||:::      :::|||   |||
Db        584 PheGlnSerGlnTyrLysSerHisPheValAlaAlaSerLeuSerAsnGlnLysAlaGly 603

Qy       2101 AAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGAT 2160
                 |||               :::|||::::::|||:::   :::   |||   |||
Db        604 SerSerAlaAlaGlyAlaSerGlyTrpThrSerAlaGlySerLeuAsnSerValProThr 623

Qy       2161 GAGTCA 2166
                 |||
Db        624 AsnSer 625
```

RESULT 9

```
US-10-195-117-2
; Sequence 2, Application US/10195117
; Publication No. US20030092083A1
; GENERAL INFORMATION:
;  APPLICANT: In2Gen Co., Ltd.
;  APPLICANT:  Jeoung, Doo-il
;  APPLICANT:  Cho, Bomsoo
;  APPLICANT:  Lim, Yoon
;  APPLICANT:  Park, Saeyoung
;  APPLICANT:  Lee, Daeyeon
;  APPLICANT:  Bang, Yung-Jue
;  APPLICANT:  Yang, Hankwang
;  APPLICANT:  Kim, Dae-Kee
;  TITLE OF INVENTION: CAGE Antigen
;  FILE REFERENCE: 59258-00002
;  CURRENT APPLICATION NUMBER: US/10/195,117
;  CURRENT FILING DATE:  2002-07-11
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 630
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

```
US-10-195-117-2

Alignment Scores:
Pred. No.:               2.89e-45       Length:         630
Score:                   622.00         Matches:        194
Percent Similarity:      46.33%         Conservative:   115
Best Local Similarity:   29.09%         Mismatches:     224
Query Match:             15.73%         Indels:         134
DB:                      9              Gaps:           24

US-09-714-865-15 (1-2172) x US-10-195-117-2 (1-630)

Qy        183 TGGGCGGAATTTTGGAAACAGAGATGCTGGTGAGTGTAATAAGCGAGAT---------AA 233
              ||||||     ||||||  |||           |||   |||   ||||||      ::
Db          4 TrpAlaProGluTrpLys-Arg------AlaGluAlaAsnProArgAspLeuGlyAlaSe 21

Qy        234 TACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGT------TTTGGAAACAGA----- 282
              :               |||       |||   |||   |||      ||||||::::::
Db         21 rTrpAspValArgGlySerArgGlySerGlyTrpSerGlyProPheGlyHisGlnGlyPr 41

Qy        282 ------------------------------------------------------------ 282

Db         41 oArgAlaAlaGlySerArgGluProProLeuCysPheLysIleLysAsnAsnMetValGl 61

Qy        283 ---------GGTTTTTCAAACAGCAGGTTTGAAGATGGTG-------------------- 313
                       |||:::|||   |||:::   :::|||   :::
Db         61 yValValIleGlyTyrSerGlySerLysIleLysAsp-LeuGlnHisSerThrAsnThrL 81

Qy        314 --ATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGCGAAGATAATCCAACACGGAA 371
                |||   :::::::::||||||||            ||||||:::          |||
Db         81 ysIleGlnIleIleAsnGlyGluSer---------GluAlaLysValArgIlePheGlyA 98

Qy        372 CAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAA---ATAATTCAGAAGCTTCAGG 428
                          :::|||          :::|||   :::|||:::|||      :::
Db         98 snArgGluMetLysAlaLysAlaLysAlaAlaIleGluThrLeuIleArgLysGlnGluS 118

Qy        429 GCCATACAGAAGAG------GTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGG 482
              :::         |||      |||:::            |||:::          :::   |||
Db        118 erTyrAsnSerGluSerSerValAspAsnAlaAlaSerGlnThrProIle--GlyArgAs 137

Qy        483 TCTAGGAAGTCCAAATAATGACTTAGACCCAGAC-------------------------- 516
                ||||||         ||||||:::      :::
Db        137 nLeuGly------ArgAsnAspIleValGlyGluAlaGluProLeuSerAsnTrpAspAr 155

Qy        517 ------------------GAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAG 557
                                ||||||   :::|||                :::
Db        155 gIleArgAlaAlaValValGluCysGluLysArgLys---------TrpAlaAspLeuPr 172

Qy        558 ACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAG 617
                ||||||                  |||                :::       ||
Db        172 oProValLysLys---------AsnPheTyrIleGluSerLysAlaThrSerCysMetSe 189

Qy        618 TGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTC 677
              ||||                               :::||||||                ::
Db        189 rGluMet-----------------------GlnValIle------------------As 195

Qy        678 TTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTA 737
              :|||:::   |||                        |||   :::          :::
Db        195 nTrpArgLysGluAsnPheAsnIleThrCysAspAspLeuLysSerGlyGluLysArgLe 215

Qy        738 CATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCAT 797
                ||||||   |||               :::|||:::
Db        215 uIleProLysProThrCysArgPheLysAspAla------------------------- 226

Qy        798 AAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGC 857
                   |||   :::|||      :::|||
Db        227 ----PheGlnGlnTyrProAspLeuLeu------------------------------- 234

Qy        858 AATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGC 917
                                                          :::|||   :::
Db        235 ------------------------------------------LysSerIleIleArgVa 240

Qy        918 TGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACG 977
```

```
                  |||     |||   ||||||:::|||       :::   |||||||||||   |||
Db        240 lGlyIleValLysProThrProIleGlnSerGlnAlaTrpProIleIleLeuGlnGlyIl 260

Qy        978 AGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAAT 1037
                  ||||||:::       ||||||||||||:::||||||||     ::::::||||:::|||
Db        260 eAspLeuIleValValAlaGlnThrGlyThrGlyLysThrLeuSerTyrLeuMetProGl 280

Qy       1038 TTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACC 1097
                     |||:::      |||          ||||||   ::::::         ||
Db        280 yPheIleHisLeu------AspSerGlnProIleSerArg---GluGlnArgAsnGlyPr 297

Qy       1098 AGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAG 1157
                  |      :::::::::   ||||||||||||||          :::       |||
Db        297 oGlyMetLeuValLeuThrProThrArgGluLeuAlaLeuHisValGluAlaGluCysSe 317

Qy       1158 AAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGG 1217
                  |||:::|||:::          ::::::::::::   |||||||||||
Db        317 rLysTyrSerTyr---LysGlyLeuLysSerIleCysIleTyrGlyGlyArgAsnArgAs 336

Qy       1218 ACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACT 1277
                     |||      |||   :::|||   :::|||:::   |||||||||||||||
Db        336 nGlyGlnIleGluAspIleSerLysGlyValAspIleIleIleAlaThrProGlyArgLe 356

Qy       1278 GATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGA 1337
                  |   |||:::                :::   |||:::   |||   ||||||||:::||
Db        356 uAsnAspLeuGlnMetAsnAsnSerValAsnLeuArgSerIleThrTyrLeuValIleAs 376

Qy       1338 TGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTG 1397
                  |||||||||:::||||||||||||   |||   |||:::::::::|||::::::
Db        376 pGluAlaAspLysMetLeuAspMetGluPheGluProGlnIleArgLysIleLeu----- 394

Qy       1398 CCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGA 1457
                         :::      :::      |||||||||:::|||   |||||||||:::|||::
Db        395 -------LeuAspValArgProAspArgGlnThrValMetThrSerAlaThrTrpProAs 412

Qy       1458 GGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGG 1517
                  :     :::::::::||||||      :::||||||   :::   :::   |||   |||||
Db        412 pThrValArgGlnLeuAlaLeuSerTyrLeuLys---AspProMetIleValTyrValGl 431

Qy       1518 ACAAGTGGGT---GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTT 1574
                  |   :::            |||         |||:::|||   ::::::         :::
Db        431 yAsnLeuAsnLeuValAlaValAsnThrValLysGlnAsnIleIleValThrThrGluLy 451

Qy       1575 CTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTAT 1631
                        ||||||         |||   :::   |||:::      :::::::::   ::
Db        451 sGluLysArgAlaLeuThrGlnGluPheValGluAsnMetSerProAsnAspLysValIl 471

Qy       1632 GGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAA 1691
                  ::::||||||      |||      ||||||      ::::::            :::
Db        471 eMetPheValSerGlnLysHisIleAlaAspAspLeuSerSerAspPheAsnIleGlnGl 491

Qy       1692 AATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGG 1751
                  ||||||      |||:::||||||:::      ||||||   ::::::|||:::|||:::
Db        491 yIleSerAlaGluSerLeuHisGlyAsnSerGluGlnSerAspGlnGluArgAlaValGl 511

Qy       1752 AGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCT 1811
                  ||||||:::   |||            :::||||||   |||   :::   :::|||||||
Db        511 uAspPheLysSerGlyAsnIleLysIleLeuIleThrThrAspIleValSerArgGlyLe 531

Qy       1812 GGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATA 1871
                  ||||:::   :::|||   ||||||   |||:::|||   |||      ||||||   ||
Db        531 uAspLeuAsnAspValThrHisValTyrAsnTyrAspPheProArgAsnIleAspValTy 551

Qy       1872 TGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTT 1931
                  |||||||||:::|||                   |||            ||||||
Db        551 rValHisArgValGlySerLeuAsp------GlyGlnGluArgLeuHisIleSerSerLe 569

Qy       1932 TGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCA 1991
                        :::   |||:::   :::|||      |||:::|||:::|||      |||
Db        569 uIleThrGlnArgAspSerLysMetAlaGlyGluLeuIleLysIleLeuAspArgAlaAs 589

Qy       1992 ACAGGATGTTCCT 2004
                  |||   ||||||
Db        589 nGlnSerValPro 593
```

RESULT 10
US-09-815-242-13491
; Sequence 13491, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13491
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13491

```
Alignment Scores:
Pred. No.:              3.26e-45      Length:         524
Score:                  621.00        Matches:        145
Percent Similarity:     57.84%        Conservative:   80
Best Local Similarity:  37.28%        Mismatches:     144
Query Match:            15.71%        Indels:         20
DB:                     10            Gaps:           7

US-09-714-865-15 (1-2172) x US-09-815-242-13491 (1-524)

Qy        862 CTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGT 921
                  :::   |||   |||   ||||||               |||         |||   |||||||||
Db          1 MetLysPheAsnGluLeuAsnLeuSerAlaAspLeuLeuAlaGluIleGluLysAlaGly 20

Qy        922 TATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGAT 981
                  :::   :::   :::|||:::|||:::      :::||||||:::   |||   |||:::|||
Db         21 PheValGluAlaSerProIleGlnGluGlnThrIleProLeuAlaLeuGluGlyLysAsp 40

Qy        982 TTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTG 1041
                  ::::::      ||||||||||||:::||||||||||||||||||   |||||||   |||
Db         41 ValIleGlyGlnAlaGlnThrGlyThrGlyLysThrAlaAlaPheGlyLeuProThrLeu 60

Qy       1042 GCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAG 1101
                        :::      :::      |||                              :::
Db         61 GluLysIleArgThrGluGluAlaThrIle---------------------------Gln 71

Qy       1102 TGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAA 1161
                  :::::::::||||||||||||||||||         |||         |||      :::
Db         72 AlaLeuValIleAlaProThrArgGluLeuAlaValGlnSerGlnGluGluLeuPheArg 91

Qy       1162 TTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACAT 1221
```

```
                   |||              |||:::          :::|||||||||:::    :::
Db         92 PheGlyArgSerLysGlyValLysValArgSerValTyrGlyGlySerSerIleGluLys 111

Qy       1222 TCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATG 1281
                 |||:::   :::     |||   :::|||:::     ||||||||||||||:::
Db        112 GlnIleLysAlaLeuLysSerGlyAlaHisIleValValGlyThrProGlyArgLeuLeu 131

Qy       1282 GATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAA 1341
              |||:::|||   ::::::  :::   |||:::   |||:::   |||:::|||||||||
Db        132 AspLeuIleLysArgLysAlaLeuLysLeuGlnAspIleGluThrLeuIleLeuAspGlu 151

Qy       1342 GCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCA 1401
              ||||||   ||||||:::|||||||||       :::::::::   :::||||||
Db        152 AlaAspGluMetLeuAsnMetGlyPheLeuGluAspIleGluAlaIleIleSer------ 169

Qy       1402 GGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAA 1461
                 :::|||      |||   ||||||||||||:::||||||||||||   |||:::
Db        170 ArgValPro------GluAsnArgGlnThrLeuLeuPheSerAlaThrMetProAspAla 187

Qy       1462 ATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCA---AATTATCTGTTTGTTGCTGTTGGA 1518
              |||:::|||:::      :::|||:::|||:::       ::::::   :::|||
Db        188 IleLysArgIleGlyValGlnPheMetLysAlaProGluHisValLysIleAlaAlaLys 207

Qy       1519 CAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCA 1578
              ::::::                  |||   |||      ::::::|||   :::
Db        208 GluLeu------ThrThrGluLeuValAspGlnTyrTyrIleArgValLysGluGlnGlu 225

Qy       1579 AAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTT 1638
              |||   :::   :::      ::::::               |||      :::||||||
Db        226 LysPheAspThrMetThrArgLeuMetAspValAlaGlnProGluLeuAlaIleValPhe 245

Qy       1639 GTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCA 1698
                       ||||||::::::   |||                  |||
Db        246 GlyArgThrLysArgArgValAspGluLeuThrArgGlyLeuLysIleArgGlyPheArg 265

Qy       1699 ACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTT 1758
                       ||||||||||||   :::|||   :::|||   :::   |||   ||||||
Db        266 AlaGluGlyIleHisGlyAspLeuAspGlnAsnLysArgLeuArgValLeuArgAspPhe 285

Qy       1759 CGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATT 1818
              :::   |||         ||||||||||||||||   ||||||||||||||||||||||||
Db        286 LysAsnGlyAsnLeuAspValLeuValAlaThrAspValAlaAlaArgGlyLeuAspIle 305

Qy       1819 GAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCAT 1878
                    |||   ||||||   |||:::|||:::|||         :::   |||||||||
Db        306 SerGlyValThrHisValTyrAsnTyrAspIleProGlnAspProGluSerTyrValHis 325

Qy       1879 CGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTT 1938
              ||||||||||||||||||||   |||      :::|||::::::|||:::|||
Db        326 ArgIleGlyArgThrGlyArgAlaGlyLysSerGlyGlnSerIleThrPheValAlaPro 345

Qy       1939 GAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGAT 1998
                       :::|||         :::::::::   ||||||       :::
Db        346 AsnGluMetGlyTyrLeu------GlnIleIleGluAsnLeuThrLysLysArgMetLys 363

Qy       1999 ------GTTCCTGCATGGTTGGAAGAA 2019
                    ||||||   :::||||||
Db        364 GlyLeuLysProAlaSerValGluGlu 372
```

RESULT 11
US-09-738-626-4778
; Sequence 4778, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4778
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4778

Alignment Scores:

Pred. No.:	6.19e-45	Length:	732
Score:	618.50	Matches:	176
Percent Similarity:	48.07%	Conservative:	98
Best Local Similarity:	30.88%	Mismatches:	208
Query Match:	15.64%	Indels:	88
DB:	9	Gaps:	16

US-09-714-865-15 (1-2172) x US-09-738-626-4778 (1-732)

```
Qy        580 AATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGT 639
              |||||||||         |||      :::         ||||||
Db          8 AsnGlyAspValGluGlnProAsnAsnValIleSerSerGlu------------------ 21

Qy        640 TTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAAT------------------------ 675
                 ::::::|||         |||      :::
Db         22 ---SerGlnGluThrProGlnGlyAspSerAlaSerAlaAspPheAlaLeuGluThrPro 40

Qy        676 ---TCTTGGAAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTG 732
                 :::      :::|||               |||   |||:::            ::::::
Db         41 ThrAsnThrValGluAspAlaProAlaSerGluGlySerGlu------------GluIle 56

Qy        733 ACCTACATACCCCCTCCTCCACCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACA 792
              |||   :::                  ||||||   ||||||      |||         :::
Db         57 ThrArgVal---AlaAspThrSerGluAspAlaAspSerAlaAspAlaAspAsnAlaSer 75

Qy        793 GGCATA---------AACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACAT 843
                 :::         |||   |||         :::                  :::
Db         76 AsnValIleAsnGluAsnGluAspSerSerGluGlyAlaAsnGlnProSerAsnGluSer 95

Qy        844 GATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAAC 903
                 :::         |||      |||:::      |||   :::   :::   :::
Db         96 SerSerThrGluAlaLysSerGlyPheAspAlaLeuGlyLeuProGluArgValLeuAsp 115

Qy        904 AACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATC 963
                 :::   |||   ||||||         :::|||:::|||      :::|||||||||
Db        116 AlaValArgLysValGlyTyrGluThrProSerProIleGlnAlaGlnThrIleProIle 135

Qy        964 ATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCT 1023
              ::::::   |||:::|||::::::         ||||||||||||:::||||||||||||||||
Db        136 LeuMetGluGlyGlnAspValValGlyLeuAlaGlnThrGlyThrGlyLysThrAlaAla 155

Qy       1024 TTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAA 1083
              |||   |||||||||||||||   :::                              |||
Db        156 PheAlaLeuProIleLeuAlaArgIle---------------------------AspLys 166

Qy       1084 GAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATT 1143
                 ::::::   |||:::   :::::::::||||||||||||||||||   |||:::
Db        167 SerValArgSerProGlnAlaLeuValLeuAlaProThrArgGluLeuAlaLeuGlnVal 186

Qy       1144 TATTTGGAAGCCAGAAAATTT---TCTTTTGGGACTTGTGTA---------AGAGCTGTT 1191
                          |||      |||   ||||||         |||               :::
Db        187 ---------AlaAspSerPheGlnSerPheAlaAspHisValGlyGlyLeuAsnValLeu 203
```

```
Qy       1192 GTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAAT 1251
                   |||||||||||          |||        :::       :::|||    :::
Db        204 ProIleTyrGlyGlyGlnAlaTyrGlyIleGlnLeuSerGlyLeuArgArgGlyAlaHis 223

Qy       1252 ATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTC 1311
              |||:::       |||||||||||::::::|||    :::    |||     :::   :::
Db        224 IleValValGlyThrProGlyArgIleIleAspHisLeuGluLysGlySerLeuAspIle 243

Qy       1312 AAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGT 1371
                 :::::::::||||||||||||||||||     ||||||:::|||||||||
Db        244 SerGlyLeuArgPheLeuValLeuAspGluAlaAspGluMetLeuAsnMetGlyPheGln 263

Qy       1372 CCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACC 1431
                 :::::::::::::::::                       :::    :::::::::|||
Db        264 GluAspValGluArgIleLeuGlu------------AspThrProAspGluLysGlnVal 279

Qy       1432 CTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAG 1491
                 :::||||||||||||     |||        |||:::||||||:::     ::::::|||
Db        280 AlaLeuPheSerAlaThrMetProAsnGlyIleArgArgLeuSerLysGlnTyrLeu--- 298

Qy       1492 TCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAG 1551
              :::|||        :::   |||                          ::::::   |||
Db        299 AsnAsnProAlaGluIleThrValLysSerGluThrArgThrAsnThrAsnIleThrGln 318

Qy       1552 ACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611
                    |||   |||          :::|||    :::   |||      |||||
Db        319 ArgPheLeuAsnValAlaHisArgAsnLysMetAspAlaLeuThrArgIleLeuGluVal 338

Qy       1612 ATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCA 1671
                       |||      ::::::||||||   ||||||    :::    :::      |||
Db        339 ThrGluPheGluAlaMetIleMetPheValArgThrLysHisGluThrGluGluValAla 358

Qy       1672 ACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGA 1731
                    |||                 |||      :::|||:::||||||       |||
Db        359 GluLysLeuArgAlaArgGlyPheSerAlaAlaAlaIleAsnGlyAspIleAlaGlnAla 378

Qy       1732 GAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACT 1791
              :::||||||:::   :::            :::   |||:::        :::||||||||||||
Db        379 GlnArgGluArgThrValAspGlnLeuLysAspGlyArgLeuAspIleLeuValAlaThr 398

Qy       1792 TCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTT 1851
                 ||||||||||||||||||||:::|||     :::    ||||||:::|||||||||:::
Db        399 AspValAlaAlaArgGlyLeuAspValGluArgIleSerHisValLeuAsnPheAspIle 418

Qy       1852 CCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACT 1911
              |||:::      :::   ||||||||||||||||||||||||||||||    |||    |||
Db        419 ProAsnAspThrGluSerTyrValHisArgIleGlyArgThrGlyArgAlaGlyArgThr 438

Qy       1912 GGCAGAGCAATTTCCTTT------------------------------------------ 1929
               |||   ||||||   |||
Db        439 GlyGluAlaIleLeuPheValThrProArgGluArgArgMetLeuArgSerIleGluArg 458

Qy       1930 -----------------------------------------TTTGATCTTGAATCGGATAAC 1950
                                                           |||      :::|||
Db        459 AlaThrAsnAlaProLeuHisGluMetGluLeuProThrValAspGlnValAsnAspPhe 478

Qy       1951 CAT---------TTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTT 2001
                          |||      :::   |||   |||   |||   |||   |||:::
Db        479 ArgLysValLysPheAlaAspSerIleThrLysSerLeuGluAspLysGlnMetAspLeu 498

Qy       2002 CCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACA 2061
                                       |||   |||   :::      :::|||
Db        499 ------------------------PheArgThrLeuValLysGluTyrSer--------- 507

Qy       2062 AGAGGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCT 2121
              :::   |||                |||              :::|||      |||
Db        508 GlnAlaAsnAspValProLeuGluAspIleAlaAlaAlaLeuAlaThrGln-AlaGlnSe 527

Qy       2122 GGGT---TTTCTTCTTCACGAGCTCCCA 2146
               |||   |||||||||   |||||||||
Db        527 rGlyAspPheLeuLeuLysGluLeuPro 536
```

RESULT 12

```
US-09-815-242-11308
; Sequence 11308, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11308
;   LENGTH: 492
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-09-815-242-11308

Alignment Scores:
Pred. No.:               1.17e-44       Length:        492
Score:                   614.50         Matches:       140
Percent Similarity:      56.62%         Conservative:  78
Best Local Similarity:   36.36%         Mismatches:    142
Query Match:             15.54%         Indels:        25
DB:                      10             Gaps:          6

US-09-714-865-15 (1-2172) x US-09-815-242-11308 (1-492)

Qy        865 ACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTAT 924
              :::|||   :::       |||   ::::::::      ::::::   :::||||||:::
Db         21 SerPheAsnAspLeuGlyLeuLysGluSerValLeuLysSerValTyrGluAlaGlyPhe 40

Qy        925 ACTAAGCTTACTCCTGTGCAAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTG 984
              |||      :::|||:::|||:::   :::||||||   :::|||   |||||||||:::
Db         41 ThrSerProSerProIleGlnGluLysAlaIleProAlaValLeuGlnGlyArgAspVal 60

Qy        985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCT 1044
              :::|||   ||||||||||||:::||||||||||||||||||   |||||||||:::
Db         61 IleAlaGlnAlaGlnThrGlyThrGlyLysThrAlaAlaPheAlaLeuProIleIle--- 79

Qy       1045 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
                                         :::      |||                |||
Db         80 --------------------------AsnAsnLeuLysAsnAsnHisThrIleGluAla 90

Qy       1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164
              :::::::::      ||||||||||||||||      ||||||      |||      |||
Db         91 LeuValIleThrProThrArgGluLeuAlaMetGlnIleSerAspGluIlePheLysLeu 110

Qy       1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA 1224
                       |||   :::   |||   :::|||||||||         :::
Db        111 GlyLysHisThrArgThrLysThrValCysValTyrGlyGlyGlnSerValLysLysGln 130
```

```
Qy       1225 ATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGAT 1284
                          |||   :::          ::::::  ||||||||||||||||||:::|||
Db        131 CysGluPheIleLysLysAsnProGlnValMetIleAlaThrProGlyArgLeuLeuAsp 150

Qy       1285 ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCT 1344
                 :::       |||:::|||               |||   :::|||||||||||:::
Db        151 HisLeuLysAsnGluArgIleHisLysPheValProLysValValValLeuAspGluSer 170

Qy       1345 GATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
                 |||   ||||||||||||||||||          :::::::::::::::
Db        171 AspGluMetLeuAspMetGlyPheLeuAspAspIleGluGluIlePheAspTyr------ 188

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
              :::||||||         :::   |||   |||:::||||||||||||    ||||||   |||
Db        189 LeuProSer------GluAlaGlnIleLeuLeuPheSerAlaThrMetProGluProIle 206

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
              :::|||||||||   :::   |||:::   |||   :::   :::   :::
Db        207 LysArgLeuAlaAspLysIleLeuGlu---AsnProIleLysIleHisIleAlaProSer 225

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
                          |||:::   |||            :::   :::         :::
Db        226 AsnIleThrAsnThrAspIleThrGlnArgPheTyrValIleAsnGluHisGluArgAla 245

Qy       1585 GAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAA 1644
              |||   ::::::   :::|||            ::::::::::::||||||
Db        246 GluAlaIleMetArgLeuLeuAspThrGlnAlaProLysLysSerIleValPheThrArg 265

Qy       1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACA 1704
              |||||||||:::||||||            ||||||      :::         :::|||
Db        266 ThrLysLysGluAlaAspGluLeuHisGlnPheLeuAlaSerLysAsnTyrLysSerThr 285

Qy       1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
              ::::::|||||||||   :::||||||:::|||         ::::::      |||:::
Db        286 AlaLeuHisGlyAspMetAspGlnArgAspArgArgSerSerIleMetAlaPheLysLys 305

Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824
                          ||||||||||||||||||   ||||||:::|||||||||||||||||||
Db        306 AsnAspAlaAspValLeuValAlaThrAspValAlaSerArgGlyLeuAspIleSerGly 325

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884
              |||   ||||||   |||:::   ||||||            :::   |||:::|||||||||
Db        326 ValSerHisValPheAsnTyrHisLeuProLeuAsnThrGluSerTyrIleHisArgIle 345

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCG 1944
              |||||||||||||||   |||      |||   ||||||:::
Db        346 GlyArgThrGlyArgAlaGlyLysLysGlyMetAlaIleThr------------------ 359

Qy       1945 GATAACCATTTAGCACAGCCTCTA---GTAAAAGTATTGACAGATGCTCAACAGGATGTT 2001
                          |||      ||||||      |||   |||         |||:::::::::
Db        360 ---------LeuValThrProLeuGluTyrLysGluLeuLeuArgMetGlnLysGluIle 376

Qy       2002 CCTGCATGGTTGGAA 2016
                 :::   :::|||
Db        377 AspSerGluIleGlu 381

RESULT 13
US-09-732-091-2
; Sequence 2, Application US/09732091
; Patent No. US20020107368A1
; GENERAL INFORMATION:
;  APPLICANT:  Tian, Jing-Hui
;  APPLICANT:  Walker, Richard I.
;  APPLICANT:  Jackson, W. James
;  TITLE OF INVENTION:  Helicobacter proteins, gene sequences and uses
;  TITLE OF INVENTION:  thereof
;  FILE REFERENCE:  7969-088
;  CURRENT APPLICATION NUMBER: US/09/732,091
;  CURRENT FILING DATE:  2000-12-07
;  NUMBER OF SEQ ID NOS:  44
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 492
;   TYPE: PRT
;   ORGANISM: Helicobacter sp.
US-09-732-091-2

Alignment Scores:
Pred. No.:              1.17e-44       Length:        492
Score:                  614.50         Matches:       140
Percent Similarity:     56.62%         Conservative:  78
Best Local Similarity:  36.36%         Mismatches:    142
Query Match:            15.54%         Indels:        25
DB:                     10             Gaps:          6

US-09-714-865-15 (1-2172) x US-09-732-091-2 (1-492)

Qy        865 ACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTAT 924
              :::|||   :::         |||   :::::::::      ::::::   :::||||||:::
Db         21 SerPheAsnAspLeuGlyLeuLysGluSerValLeuLysSerValTyrGluAlaGlyPhe 40

Qy        925 ACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTG 984
              |||      :::|||:::|||:::   :::||||||   :::|||   |||||||||:::
Db         41 ThrSerProSerProIleGlnGluLysAlaIleProAlaValLeuGlnGlyArgAspVal 60

Qy        985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTGGCT 1044
              :::|||   ||||||||||||:::||||||||||||||||||   |||||||||:::
Db         61 IleAlaGlnAlaGlnThrGlyThrGlyLysThrAlaAlaPheAlaLeuProIleIle--- 79

Qy       1045 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
                                          :::   |||               |||
Db         80 --------------------------AsnAsnLeuLysAsnAsnHisThrIleGluAla 90

Qy       1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164
              :::::::::   |||||||||||||||      ||||||      |||      |||
Db         91 LeuValIleThrProThrArgGluLeuAlaMetGlnIleSerAspGluIlePheLysLeu 110

Qy       1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA 1224
                          |||      :::   |||   :::|||||||||      :::
Db        111 GlyLysHisThrArgThrLysThrValCysValTyrGlyGlyGlnSerValLysLysGln 130

Qy       1225 ATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGAT 1284
                          |||   :::         ::::::   ||||||||||||||||||:::|||
Db        131 CysGluPheIleLysLysAsnProGlnValMetIleAlaThrProGlyArgLeuLeuAsp 150

Qy       1285 ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCT 1344
                 :::       |||:::|||               |||   :::|||||||||||:::
Db        151 HisLeuLysAsnGluArgIleHisLysPheValProLysValValValLeuAspGluSer 170

Qy       1345 GATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
              |||   ||||||||||||||||||          :::::::::::::::
Db        171 AspGluMetLeuAspMetGlyPheLeuAspAspIleGluGluIlePheAspTyr------ 188

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
              :::||||||         :::   |||   |||:::||||||||||||   ||||||   |||
Db        189 LeuProSer------GluAlaGlnIleLeuLeuPheSerAlaThrMetProGluProIle 206

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
              :::|||||||||   :::   |||:::   |||   :::   :::   :::
Db        207 LysArgLeuAlaAspLysIleLeuGlu---AsnProIleLysIleHisIleAlaProSer 225

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
                          |||:::   |||            :::   :::         :::
Db        226 AsnIleThrAsnThrAspIleThrGlnArgPheTyrValIleAsnGluHisGluArgAla 245

Qy       1585 GAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAA 1644
              |||   ::::::   :::|||            ::::::::::::||||||
Db        246 GluAlaIleMetArgLeuLeuAspThrGlnAlaProLysLysSerIleValPheThrArg 265

Qy       1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACA 1704
              |||||||||:::||||||            ||||||      :::         :::|||
Db        266 ThrLysLysGluAlaAspGluLeuHisGlnPheLeuAlaSerLysAsnTyrLysSerThr 285

Qy       1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
              ::::::|||||||||   :::||||||:::|||         ::::::      |||:::
Db        286 AlaLeuHisGlyAspMetAspGlnArgAspArgArgSerSerIleMetAlaPheLysLys 305
```

```
Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824
                           ||||||||||||||    ||||||:::|||||||||||||
Db        306 AsnAspAlaAspValLeuValAlaThrAspValAlaSerArgGlyLeuAspIleSerGly 325

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884
              |||  ||||||  |||:::  ||||||        :::   |||:::||||||||
Db        326 ValSerHisValPheAsnTyrHisLeuProLeuAsnThrGluSerTyrIleHisArgIle 345

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCG 1944
              ||||||||||||||  |||    |||  ||||||:::
Db        346 GlyArgThrGlyArgAlaGlyLysLysGlyMetAlaIleThr----------------- 359

Qy       1945 GATAACCATTTAGCACAGCCTCTA---GTAAAGTATTGACAGATGCTCAACAGGATGTT 2001
                        |||   ||||||   |||   |||        |||:::::::::
Db        360 ---------LeuValThrProLeuGluTyrLysGluLeuLeuArgMetGlnLysGluIle 376

Qy       2002 CCTGCATGGTTGGAA 2016
                 :::   :::|||
Db        377 AspSerGluIleGlu 381
```

RESULT 14
US-09-732-091-42
; Sequence 42, Application US/09732091
; Patent No. US20020107368A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
;   LENGTH: 504
;   TYPE: PRT
;   ORGANISM: Helicobacter sp.
US-09-732-091-42

Alignment Scores:

Pred. No.:	1.18e-44	Length:	504
Score:	614.50	Matches:	140
Percent Similarity:	56.62%	Conservative:	78
Best Local Similarity:	36.36%	Mismatches:	142
Query Match:	15.54%	Indels:	25
DB:	10	Gaps:	6

US-09-714-865-15 (1-2172) x US-09-732-091-42 (1-504)

```
Qy        865 ACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTAT 924
              :::|||   :::      |||   :::::::::    ::::::   :::|||||||:::
Db         33 SerPheAsnAspLeuGlyLeuLysGluSerValLeuLysSerValTyrGluAlaGlyPhe 52

Qy        925 ACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTG 984
              |||     :::|||:::|||:::   :::||||||   :::|||   |||||||||:::
Db         53 ThrSerProSerProIleGlnGluLysAlaIleProAlaValLeuGlnGlyArgAspVal 72

Qy        985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCT 1044
              :::|||   |||||||||||:::||||||||||||||||   |||||||||:::
Db         73 IleAlaGlnAlaGlnThrGlyThrGlyLysThrAlaAlaPheAlaLeuProIleIle--- 91

Qy       1045 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
                                         :::      |||               |||
Db         92 ---------------------------AsnAsnLeuLysAsnAsnHisThrIleGluAla 102

Qy       1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164
              :::::::::   |||||||||||||||      ||||||      |||       |||
Db        103 LeuValIleThrProThrArgGluLeuAlaMetGlnIleSerAspGluIlePheLysLeu 122

Qy       1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA 1224
                       |||      :::   |||    :::|||||||||       :::
Db        123 GlyLysHisThrArgThrLysThrValCysValTyrGlyGlyGlnSerValLysLysGln 142

Qy       1225 ATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGAT 1284
                 |||   :::       ::::::   ||||||||||||||||:::|||
Db        143 CysGluPheIleLysLysAsnProGlnValMetIleAlaThrProGlyArgLeuLeuAsp 162

Qy       1285 ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCT 1344
                 :::         |||:::|||         |||   :::|||||||||||||:::
Db        163 HisLeuLysAsnGluArgIleHisLysPheValProLysValValValLeuAspGluSer 182

Qy       1345 GATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGA 1404
              |||   ||||||||||||||||      :::::::::::::::
Db        183 AspGluMetLeuAspMetGlyPheLeuAspAspIleGluGluIlePheAspTyr------ 200

Qy       1405 ATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATT 1464
              :::||||||      :::   |||   |||:::||||||||||||   ||||||   |||
Db        201 LeuProSer------GluAlaGlnIleLeuLeuPheSerAlaThrMetProGluProIle 218

Qy       1465 CAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTG 1524
              :::|||||||||   :::   |||:::   |||   :::   :::   :::
Db        219 LysArgLeuAlaAspLysIleLeuGlu---AsnProIleLysIleHisIleAlaProSer 237

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
                          |||:::   |||             :::   :::      :::
Db        238 AsnIleThrAsnThrAspIleThrGlnArgPheTyrValIleAsnGluHisGluArgAla 257

Qy       1585 GAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAA 1644
              |||   ::::::   :::|||            ::::::::::::|||||
Db        258 GluAlaIleMetArgLeuLeuAspThrGlnAlaProLysLysSerIleValPheThrArg 277

Qy       1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACA 1704
              |||||||||:::||||||            ||||||      :::         :::|||
Db        278 ThrLysLysGluAlaAspGluLeuHisGlnPheLeuAlaSerLysAsnTyrLysSerThr 297

Qy       1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
              ::::::|||||||||   :::||||||:::|||      ::::::      |||:::
Db        298 AlaLeuHisGlyAspMetAspGlnArgAspArgArgSerSerIleMetAlaPheLysLys 317

Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824
                           ||||||||||||||    ||||||:::|||||||||||||
Db        318 AsnAspAlaAspValLeuValAlaThrAspValAlaSerArgGlyLeuAspIleSerGly 337

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884
              |||  ||||||  |||:::  ||||||        :::   |||:::||||||||
Db        338 ValSerHisValPheAsnTyrHisLeuProLeuAsnThrGluSerTyrIleHisArgIle 357

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCG 1944
              ||||||||||||||  |||    |||  ||||||:::
Db        358 GlyArgThrGlyArgAlaGlyLysLysGlyMetAlaIleThr----------------- 371

Qy       1945 GATAACCATTTAGCACAGCCTCTA---GTAAAGTATTGACAGATGCTCAACAGGATGTT 2001
                        |||   ||||||   |||   |||        |||:::::::::
Db        372 ---------LeuValThrProLeuGluTyrLysGluLeuLeuArgMetGlnLysGluIle 388

Qy       2002 CCTGCATGGTTGGAA 2016
                 :::   :::|||
Db        389 AspSerGluIleGlu 393
```

RESULT 15
US-09-963-790A-2
; Sequence 2, Application US/09963790A
; Patent No. US20020115161A1
; GENERAL INFORMATION:
; APPLICANT: FARWICK, Mike, et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE deaD GENE
; FILE REFERENCE: 032301 WD 230
; CURRENT APPLICATION NUMBER: US/09/963,790A
; CURRENT FILING DATE: 2001-09-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 624

```
;    TYPE: PRT
;    ORGANISM: Corynebacterium glutamicum
US-09-963-790A-2

Alignment Scores:
Pred. No.:              1.74e-43      Length:        624
Score:                  601.50        Matches:       153
Percent Similarity:     51.67%        Conservative:  79
Best Local Similarity:  34.08%        Mismatches:    153
Query Match:            15.21%        Indels:        64
DB:                     10            Gaps:          11

US-09-714-865-15 (1-2172) x US-09-963-790A-2 (1-624)

Qy        907 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 966
              :::    |||   ||||||         :::|||:::|||      :::||||||||:::
Db          9 ValArgLysValGlyTyrGluThrProSerProIleGlnAlaGlnThrIleProIleLeu 28

Qy        967 CTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              :::    |||:::|||::::::         ||||||||||:::||||||||||||||||
Db         29 MetGluGlyGlnAspValValGlyLeuAlaGlnThrGlyThrGlyLysThrAlaAlaPhe 48

Qy       1027 CTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG 1086
                 ||||||||||||||   :::                            |||
Db         49 AlaLeuProIleLeuAlaArgIle---------------------------AspLysSer 59

Qy       1087 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
              ::::::   |||:::   :::::::::|||||||||||||||       |||:::
Db         60 ValArgSerProGlnAlaLeuValLeuAlaProThrArgGluGlnAlaLeuGlnVal--- 78

Qy       1147 TTGGAAGCCAGAAAATTT---TCTTTTGGGACTTGTGTA---------AGAGCTGTTGTT 1194
                    |||     |||   ||||||      |||               :::
Db         79 ------AlaAspSerPheGlnSerPheAlaAspHisValGlyGlyLeuAsnValLeuPro 96

Qy       1195 ATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATA 1254
              ||||||||||||         |||      :::      :::   :::|||   :::|||
Db         97 IleTyrGlyGlyGlnAlaTyrGlyIleGlnLeuSerGlyLeuArgArgGlyAlaHisIle 116

Qy       1255 TTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAA 1314
              :::      ||||||||||||::::::|||   :::   |||      :::   :::
Db        117 ValValGlyThrProGlyArgIleIleAspHisLeuGluLysGlySerLeuAspIleSer 136

Qy       1315 CAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCA 1374
                 :::::::::||||||||||||||||||||   ||||||:::|||||||||
Db        137 GlyLeuArgPheLeuValLeuAspGluAlaAspGluMetLeuAsnMetGlyPheGlnGlu 156

Qy       1375 GAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTT 1434
              ::::::::::::::::::                     :::   :::::::::|||
Db        157 AspValGluArgIleLeuGlu------------AspThrProAspGluLysGlnValAla 172

Qy       1435 ATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCA 1494
              :::||||||||||||   |||      |||:::||||||:::   ::::::|||   :::
Db        173 LeuPheSerAlaThrMetProAsnGlyIleArgArgLeuSerLysGlnTyrLeu---Asn 191

Qy       1495 AATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACC 1554
              |||         :::   |||                     ::::::   |||
Db        192 AsnProAlaGluIleThrValLysSerGluThrArgThrAsnThrAsnIleThrGlnArg 211

Qy       1555 GTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATA 1614
                 |||   |||      :::|||   :::   |||      ||||||
Db        212 PheLeuAsnValAlaHisArgAsnLysMetAspAlaLeuThrArgIleLeuGluValThr 231

Qy       1615 GGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACT 1674
                 |||      ::::::||||||   ||||||   :::   :::      |||
Db        232 GluPheGluAlaMetIleMetPheValArgThrLysHisGluThrGluGluValAlaGlu 251

Qy       1675 TTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAG 1734
                 |||         |||      :::|||:::||||||      |||   :::
Db        252 LysLeuArgAlaArgGlyPheSerAlaAlaAlaIleAsnGlyAspIleAlaGlnAlaGln 271

Qy       1735 CGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCA 1794
              ||||||:::   :::         :::   |||:::      :::||||||||||||
Db        272 ArgGluArgThrValAspGlnLeuLysAspGlyArgLeuAspIleLeuValAlaThrAsp 291

Qy       1795 GTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCT 1854
              ||||||||||||||||||||:::|||      :::      ||||||:::|||||||||:::|||
Db        292 ValAlaAlaArgGlyLeuAspValGluArgIleSerHisValLeuAsnPheAspIlePro 311

Qy       1855 TCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGC 1914
              :::      :::   |||||||||||||||||||||||||||      |||   ||||||
Db        312 AsnAspThrGluSerTyrValHisArgIleGlyArgThrGlyArgAlaGlyArgThrGly 331

Qy       1915 AGAGCAATTTCCTTT--------------------------------------------- 1929
                 ||||||   |||
Db        332 GluAlaIleLeuPheValThrProArgGluArgArgMetLeuArgSerIleGluArgAla 351

Qy       1930 ------------------------------------TTTGATCTTGAATCGGATAACCAT 1953
                                                  |||      :::|||
Db        352 ThrAsnAlaProLeuHisGluMetGluLeuProThrValAspGlnValAsnAspPheArg 371

Qy       1954 ---------TTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCT 2004
                       |||      :::   |||   |||   |||   |||   |||:::
Db        372 LysValLysPheAlaAspSerIleThrLysSerLeuGluAspLysGlnMetAspLeu--- 390

Qy       2005 GCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAGTACAAGA 2064
                                   |||   |||   :::      :::|||         :::
Db        391 --------------------PheArgThrLeuValLysGluTyrSer---------Gln 400

Qy       2065 GGAAACGTGTTTGCATCAGTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGG 2124
                 |||               |||               :::|||      |||   ||
Db        401 AlaAsnAspValProLeuGluAspIleAlaAlaAlaLeuAlaThrGln-AlaGlnSerGl 420

Qy       2125 T---TTTCTTCTTCACGAGCTCCCA 2146
              |   |||||||||   |||||||||
Db        420 yAspPheLeuLeuLysGluLeuPro 428

Search completed: June 10, 2003, 17:02:55
Job time : 186.5 secs
```

OM nucleic - protein search, using frame_plus_n2p model

```
Run on:         June 10, 2003, 16:35:50 ; Search time 24 Seconds
                                          (without alignments)
                                          5325.549 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  3954
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:      525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09714865/runat_05062003_111759_26063/app_query.fasta_1.2311
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09714865_@CGN_1_1_16_@runat_05062003_111759_26063 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1215	30.7	662	3	US-09-058-489-15	Sequence 15, Appl
2	1215	30.7	662	3	US-09-058-489-16	Sequence 16, Appl
3	1182.5	29.9	660	3	US-09-058-489-18	Sequence 18, Appl
4	1182.5	29.9	660	3	US-09-058-489-91	Sequence 91, Appl
5	709.5	17.9	648	4	US-09-183-706-43	Sequence 43, Appl
6	709.5	17.9	648	4	US-09-567-995-43	Sequence 43, Appl
7	704	17.8	1261	4	US-09-208-742-4	Sequence 4, Appli
8	704	17.8	1261	4	US-09-332-295-2	Sequence 2, Appli
9	704	17.8	1261	4	US-09-709-979-2	Sequence 2, Appli
10	566.5	14.3	528	4	US-09-134-001C-4262	Sequence 4262, Ap
11	505.5	12.8	411	4	US-09-318-443-8	Sequence 8, Appli
12	504.5	12.8	415	4	US-09-318-443-2	Sequence 2, Appli
13	503.5	12.7	411	4	US-09-318-443-6	Sequence 6, Appli
14	488	12.3	403	2	US-08-607-509-4	Sequence 4, Appli
15	488	12.3	403	2	US-08-634-642-4	Sequence 4, Appli
16	488	12.3	403	3	US-08-989-370-4	Sequence 4, Appli
17	485	12.3	403	2	US-08-533-669A-10	Sequence 10, Appl
18	485	12.3	403	2	US-08-607-509-2	Sequence 2, Appli
19	485	12.3	403	2	US-08-454-036-2	Sequence 2, Appli
20	485	12.3	403	2	US-08-634-642-2	Sequence 2, Appli
21	485	12.3	403	3	US-08-989-370-2	Sequence 2, Appli
22	485	12.3	403	4	US-09-183-861-10	Sequence 10, Appl
23	485	12.3	403	4	US-09-022-765-10	Sequence 10, Appl
24	485	12.3	403	5	PCT-US95-05064-2	Sequence 2, Appli
25	473	12.0	455	4	US-09-134-001C-3653	Sequence 3653, Ap
26	469.5	11.9	407	3	US-08-989-370-6	Sequence 6, Appli
27	467.5	11.8	407	3	US-08-989-370-5	Sequence 5, Appli
28	465.5	11.8	491	4	US-09-314-268-3	Sequence 3, Appli
29	463.5	11.7	448	2	US-08-959-749-2	Sequence 2, Appli
30	463.5	11.7	448	4	US-09-351-497-2	Sequence 2, Appli
31	416	10.5	746	4	US-09-149-934-4	Sequence 4, Appli
32	384	9.7	551	4	US-09-503-391-2	Sequence 2, Appli
33	384	9.7	551	4	US-09-503-391-4	Sequence 4, Appli
34	374	9.5	547	3	US-09-176-657-1	Sequence 1, Appli
35	342.5	8.7	859	4	US-09-149-934-1	Sequence 1, Appli
36	334.5	8.5	560	3	US-09-176-657-7	Sequence 7, Appli
37	284.5	7.2	479	4	US-09-149-934-3	Sequence 3, Appli
38	271	6.9	149	4	US-08-679-493A-85	Sequence 85, Appl
39	244.5	6.2	177	4	US-09-318-443-4	Sequence 4, Appli
40	231.5	5.9	370	2	US-08-559-303B-77	Sequence 77, Appl
41	231.5	5.9	370	4	US-09-175-828-77	Sequence 77, Appl
42	231.5	5.9	607	3	US-08-781-891-75	Sequence 75, Appl
43	230.5	5.8	334	4	US-09-463-702A-38	Sequence 38, Appl
44	223	5.6	637	4	US-09-134-001C-5658	Sequence 5658, Ap
45	221.5	5.6	459	4	US-09-134-001C-3178	Sequence 3178, Ap

ALIGNMENTS

```
RESULT 1
US-09-058-489-15
; Sequence 15, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 15
;   LENGTH: 662
;   TYPE: PRT
;   ORGANISM: Human
US-09-058-489-15

Alignment Scores:
Pred. No.:               2.36e-118     Length:         662
Score:                   1215.00       Matches:        293
Percent Similarity:      56.64%        Conservative:   91
Best Local Similarity:   43.22%        Mismatches:     180
Query Match:             30.73%        Indels:         114
DB:                      3             Gaps:           20

US-09-714-865-15 (1-2172) x US-09-058-489-15 (1-662)

Qy        214 GAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTTGGAGTTGGAAAGAGT--- 270
              :::    |||      ||||||   |||       ||||||             |||
Db         20 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 37
```

```
Qy        271 ---------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
                             ||||||      :::           |||   |||   |||
Db         38 TyrIleProProHisLeuArgAsnArgGluAlaThrArgGly---PheTyrAspLysAsp 56

Qy        316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------------------GAA 354
              |||||||||:::                   |||
Db         57 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76

Qy        355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
              |||:::      ::::::     ||||||  ||||||
Db         77 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer------------------ 90

Qy        415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGC 468
                                        |||   ||||||   |||     ||||||
Db         91 ---------------------------GlySerArgGlyArgPheAspAspArgGlyArg 101

Qy        469 CGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG 528
                            |||:::||||||
Db        102 SerAspTyrAspGlyIleGlySer------------------------------------ 109

Qy        529 CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGAT 588
                                                                  ||||||
Db        110 ---------------------------------------------------ArgGlyAsp 112

Qy        589 ACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAAT 645
                          ||||||||  |||      ||||||||||||            |||
Db        113 ------------ArgSerGlyPheGlyLysPheGluArgGlyGly------------Asn 124

Qy        646 GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
                           |||   ::::::   |||
Db        125 SerArgTrpCysAspLysSerAspGluAspAspTrpSerLys------------------ 138

Qy        706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--- 762
                                                  |||   ||||||  |||
Db        139 ------------------------------------ProLeuProProSerGluArgLeu 146

Qy        763 GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATT 822
              |||      :::|||:::              |||||||||||||:::|||||||||   |||
Db        147 GluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIle 166

Qy        823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                 ||||||   :::|||::::::   ||||||   |||   :::|||   :::
Db        167 ProValGluAlaThrGlyAsnAsnCysProProHisIleGluSerPheSerAspValGlu 186

Qy        883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
              :::   :::   :::         ||||||            ||||||:::   |||||||||
Db        187 MetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgProThrProVal 206

Qy        943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
              ||||||::::::||||||||||||         ||||||||||||||||||||||||||||||
Db        207 GlnLysHisAlaIleProIleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThr 226

Qy       1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA 1062
              ||||||||||||||||||||||||||||||||||||||||||:::      :::      ||||||
Db        227 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 246

Qy       1063 ------------ATAACTGCCAGTCGT----------------TTTAAAGAGTTGCAGGAA 1095
                          :::   |||   :::                   :::      |||
Db        247 ProGlyGluAlaLeuArgAlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyr 266

Qy       1096 CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCC 1155
              |||         :::::::::||||||||||||||||||         |||||||||   ||||||
Db        267 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluGluAla 286

Qy       1156 AGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTG 1215
              ||||||||||||:::   :::   ||||||      |||:::|||||||||         :::
Db        287 ArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGlyAlaAspIle 306

Qy       1216 GGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGA 1275
              |||      ||||||   :::   :::||||||::::::|||   ||||||||||||||||||
Db        307 GlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArg 326

Qy       1276 CTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTG 1335
              |||:::|||::::::   :::   ||||||||||||            ||||||||||||||||
Db        327 LeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeu 346

Qy       1336 GATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT 1395
              ||||||||||||||||||||||||||||||||||||   |||:::::::::::::::::::
Db        347 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArgIleValGlu 366

Qy       1396 TGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCA 1455
                       ||||||   |||         |||   |||:::||||||||||||||||||||
Db        367 GlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAlaThrPhePro 386

Qy       1456 GAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT 1515
              :::|||||||||   ||||||   :::||||||            |||:::|||:::||||||
Db        387 LysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePheLeuAlaVal 405

Qy       1516 GGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTC 1575
              |||:::||||||            ::::::   |||   |||:::   |||      :::
Db        406 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpValGluGluSer 425

Qy       1576 TCAAAAAGAGAAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATG 1632
                 ||||||         |||:::::::::|||      |||   |||      |||:::
Db        426 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeu 445

Qy       1633 GTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAA 1692
              ||||||||||||||||||||   ||||||               ||||||      |||
Db        446 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGly 465

Qy       1693 ATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGA 1752
                 :::   ||||||||||||||||||||   ||||||:::||||||:::||||||
Db        466 TyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHis 485

Qy       1753 GATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTG 1812
                 ||||||   ||||||   |||:::||||||||||||:::||||||||||||||||||
Db        486 GlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 505

Qy       1813 GATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATAT 1872
              ||||||   ||||||:::||||||||||||||||||||||||||||||   |||:::||||||
Db        506 AspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIleGluGluTyr 525

Qy       1873 GTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTT 1932
              ||||||||||||||||||||||||   ||||||   |||   |||   |||||||||
Db        526 ValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 545

Qy       1933 GATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAA 1992
              :::   |||   :::   ::::::   :::   |||:::   :::|||   :::|||:::
Db        546 Asn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuValGluAlaLys 564

Qy       1993 CAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGT 2052
              |||:::||||||:::|||||||||   :::|||:::         :::         :::
Db        565 GlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis---------TyrLys 581

Qy       2053 GGTAGTACAAGAGGAAACGTG-----------------------TTTGCATCAGTTGAT 2088
              ||||||:::||||||                             |||   :::   |||
Db        582 GlySerSerArgGlyArgSerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAsp 601

Qy       2089 ACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                 |||:::      |||      :::::::::   ||||||||||||||||||
Db        602 TyrArgGln---SerSerGlyAlaSerSerSerSerPheSerSerSerArgAla 618
```

RESULT 2
US-09-058-489-16
; Sequence 16, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; TITLE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10

```
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
;   LENGTH: 662
;   TYPE: PRT
;   ORGANISM: Human
US-09-058-489-16

Alignment Scores:
Pred. No.:               2.36e-118      Length:        662
Score:                   1215.00        Matches:       293
Percent Similarity:      56.64%         Conservative:  91
Best Local Similarity:   43.22%         Mismatches:    180
Query Match:             30.73%         Indels:        114
DB:                      3              Gaps:          20

US-09-714-865-15 (1-2172) x US-09-058-489-16 (1-662)

Qy      214 GAGTGTAATAAGCGAGATAATACATCCACAATGGGTGGTTTGGAGTTGGAAAGAGT--- 270
            :::   |||     ||||||  |||      ||||||       |||
Db       20 AspLeuAsnSerSerAspAsnGlnSer------GlyGlySerThrAlaSerLysGlyArg 37

Qy      271 ---------------TTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGAT 315
                           ||||||       :::        |||   |||   |||
Db       38 TyrIleProProHisLeuArgAsnArgGluAlaThrArgGly---PheTyrAspLysAsp 56

Qy      316 AGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC---------------------GAA 354
            |||||||||:::                   |||
Db       57 SerSerGlyTrpSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76

Qy      355 GATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAAT 414
            |||:::      ::::::      ||||||    ||||||
Db       77 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer------------------ 90

Qy      415 TCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC------CGAGGTTGC 468
                                      |||   ||||||   |||        ||||||
Db       91 --------------------------GlySerArgGlyArgPheAspAspArgGlyArg 101

Qy      469 CGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAG 528
                       |||:::||||||
Db      102 SerAspTyrAspGlyIleGlySer------------------------------------ 109

Qy      529 CGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGAT 588
                                                                  ||||||
Db      110 ---------------------------------------------------ArgGlyAsp 112

Qy      589 ACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTACAAAGGTTTAAAT 645
                        |||||||||  |||      |||||||||||            |||
Db      113 ------------ArgSerGlyPheGlyLysPheGluArgGlyGly------------Asn 124

Qy      646 GAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCAGAAGGAGGAGAA 705
                                |||   ::::::   |||
Db      125 SerArgTrpCysAspLysSerAspGluAspAspTrpSerLys------------------ 138

Qy      706 AGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--- 762
                                                   |||   ||||||   |||
Db      139 ------------------------------------ProLeuProProSerGluArgLeu 146

Qy      763 GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATT 822
            |.||      :::|||:::             ||||||||||||||:::|||||||||   |||
Db      147 GluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIle 166

Qy      823 CTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAAT 882
                ||||||   :::|||::::::   ||||||   |||   :::|||   :::
Db      167 ProValGluAlaThrGlyAsnAsnCysProProHisIleGluSerPheSerAspValGlu 186

Qy      883 CTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTG 942
            :::   :::   :::            ||||||         |||||||:::   |||||||||
Db      187 MetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgProThrProVal 206

Qy      943 CAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACA 1002
            ||||||::::::||||||||||||            ||||||||||||||||||||||||
Db      207 GlnLysHisAlaIleProIleIleLysGluLysArgAspLeuMetAlaCysAlaGlnThr 226

Qy     1003 GGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGA 1062
            |||||||||||||||||||||||||||||||||||||:::   :::      ||||||
Db      227 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 246

Qy     1063 ------------ATAACTGCCAGTCGT---------------TTTAAAGAGTTGCAGGAA 1095
                        :::   |||   :::               :::         |||
Db      247 ProGlyGluAlaLeuArgAlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyr 266

Qy     1096 CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCC 1155
            |||      :::::::::||||||||||||||||||     |||||||||   ||||||
Db      267 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluGluAla 286

Qy     1156 AGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTG 1215
            ||||||||||||:::   :::   ||||||      |||:::|||||||||      :::
Db      287 ArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGlyAlaAspIle 306

Qy     1216 GGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGA 1275
            |||      ||||||   :::   :::||||||::::::|||   |||||||||||||||
Db      307 GlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArg 326

Qy     1276 CTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTG 1335
            |||:::|||::::::   :::   ||||||||||||      ||||||||||||||||
Db      327 LeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeu 346

Qy     1336 GATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT 1395
            |||||||||||||||||||||||||||||||||   |||:::::::::::::::::
Db      347 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArgIleValGlu 366

Qy     1396 TGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCA 1455
                        ||||||   |||      |||   |||:::||||||||||||||||||
Db      367 GlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAlaThrPhePro 386

Qy     1456 GAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT 1515
            :::|||||||||   ||||||   :::||||||         |||:::|||:::||||||
Db      387 LysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePheLeuAlaVal 405

Qy     1516 GGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTC 1575
            |||:::||||||         ::::::   |||   |||:::   |||      :::
Db      406 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValValTrpValGluGluSer 425

Qy     1576 TCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG---GATGAAAGAACTATG 1632
               ||||||      |||::::::::|||      |||   |||      |||:::
Db      426 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeu 445

Qy     1633 GTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAA 1692
            ||||||||||||||||||||||   ||||||            ||||||      |||
Db      446 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGly 465

Qy     1693 ATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGA 1752
               :::   ||||||||||||||||||||||   ||||||:::||||||:::||||||
Db      466 TyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHis 485

Qy     1753 GATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTG 1812
               ||||||   ||||||   |||:::||||||||||||:::||||||||||||||||||
Db      486 GlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 505

Qy     1813 GATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATAT 1872
            ||||||   |||||||||:::||||||||||||||||||||||||   |||:::||||||
Db      506 AspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIleGluGluTyr 525

Qy     1873 GTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTT 1932
            ||||||||||||||||||||||||   ||||||   |||   |||   |||||||||
Db      526 ValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 545

Qy     1933 GATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAA 1992
            :::   |||   :::   ::::::   :::   |||:::   :::|||   :::|||:::
Db      546 Asn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuLeuValGluAlaLys 564

Qy     1993 CAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGT 2052
            |||:::||||||:::|||||||||   :::|||:::          :::          :::
```

```
Db        565 GlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis---------TyrLys 581

Qy       2053 GGTAGTACAAGAGGAAACGTG-----------------------TTTGCATCAGTTGAT 2088
              ||||||:::||||||                        |||   :::   |||
Db        582 GlySerSerArgGlyArgSerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAsp 601

Qy       2089 ACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCT 2142
                 |||:::      |||       :::::::::   ||||||||||||||||
Db        602 TyrArgGln---SerSerGlyAlaSerSerSerSerPheSerSerSerArgAla 618

RESULT 3
US-09-058-489-18
; Sequence 18, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
;   LENGTH: 660
;   TYPE: PRT
;   ORGANISM: Human
US-09-058-489-18

Alignment Scores:
Pred. No.:              6.01e-115    Length:        660
Score:                  1182.50      Matches:       278
Percent Similarity:     58.06%       Conservative:  93
Best Local Similarity:  43.51%       Mismatches:    167
Query Match:            29.91%       Indels:        101
DB:                     3            Gaps:          18

US-09-714-865-15 (1-2172) x US-09-058-489-18 (1-660)

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------- 351
              |||   |||   |||||||||||:::                  |||
Db         52 PheHisAspLysAspSerSerGlyTrpSerCysSerLysAspLysAspAlaTyrSerSer 71

Qy        352 ---------GAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
                       |||:::      :::        ||||||:::||||||
Db         72 PheGlySerArgAspSerArgGlyLysProGlyTyrPheSerGluArgGlySer------ 89

Qy        403 GATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC--- 459
                                                    |||   ||||||   |||
Db         90 ---------------------------------------GlySerArgGlyArgPheAsp 96

Qy        460 ---CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGAC 516
                 ||||||                             :::|||   |||
Db         97 AspArgGly-----------------------------ArgSerAspTyrAsp------ 104

Qy        517 GAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACA 576
                                                                |||
Db        105 ------------------------------------------------------GlyIle 106

Qy        577 GGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTAC 633
              ||||||           :::   |||   |||   |||      ||||||   |||:::
Db        107 GlyAsn------------ArgGluArgProGlyPheGlyArgPheGluArgSerGlyHis 122

Qy        634 AAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCA 693
                          ::::::             |||   ::::::   |||
Db        123 SerArgTrpCysAspLys------------SerValGluAspAspTrpSerLys------ 136

Qy        694 GAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCA 753
                                                          |||   ||||||
Db        137 ------------------------------------------------ProLeuProPro 140

Qy        754 CCTGAGGAT---GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAA 810
                 |||      |||         :::|||:::         ||||||||||||||:::|||
Db        141 SerGluArgLeuGluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLys 160

Qy        811 TACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTT 870
              ||||||   |||   ||||||   :::|||   :::   ||||||   |||.     |||
Db        161 TyrAspAspIleProValGluAlaThrGlySerAsnCysProProHisIleGluAsnPhe 180

Qy        871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
                 :::   ::::::   :::   :::         ||||||         ||||||:::
Db        181 SerAspIleAspMetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArg 200

Qy        931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
                 ||||||||||||||:::::::||||||||||||         |||||||||:::|||
Db        201 ProThrProValGlnLysHisAlaIleProIleIleLysGlyLysArgAspLeuValAla 220

Qy        991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTGGCTCATATG 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::   :::
Db        221 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIle 240

Qy       1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA--------------- 1095
                    ||||||                 |||   ::::::|||
Db        241 TyrThrAspGly---ProGlyGluAlaLeuLysAlaValLysGluAsnGlyArgTyrGly 259

Qy       1096 ---------------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
                             |||      :::::::::||||||||||||||||||   |||
Db        260 ArgArgLysGlnTyrProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGln 279

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
              ||||||   ||||||||||||||||||:::   :::      ||||||      |||:::|||
Db        280 IleTyrGluGluAlaArgLysPheSerTyrArgSerArgValArgProCysValValTyr 299

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
              ||||||      :::|||      ||||||   :::   :::||||||::::::|||
Db        300 GlyGlyAlaAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuVal 319

Qy       1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
              ||||||||||||||||||:::|||::::::   :::   ||||||||||||
Db        320 AlaThrProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCys 339

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
              |||||||||||||||||||||||||||||||||||||||||||||||   |||::::::
Db        340 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIle 359

Qy       1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
              ::::::::::::      ||||||      |||      |||   |||:::||||||
Db        360 ArgArgIleValGluGlnAspThrMetProProLysGlyValArgHisThrMetMetPhe 379

Qy       1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
              ||||||||||||||:::|||||||||   ||||||   :::||||||         |||
Db        380 SerAlaThrPheProLysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyr 398

Qy       1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
              :::|||:::|||||||||:::||||||            ::::::   |||   |||:::
Db        399 IlePheLeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValVal 418

Qy       1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG--- 1617
                 |||         ||||||      |||::::::||||||         |||
Db        419 TrpValGluAspLeuAspLysArgSerPheLeuLeuAspIleLeuGlyAlaThrGlySer 438

Qy       1618 GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTT 1677
              |||      |||:::||||||||||||||||||||   ||||||         |||
Db        439 AspSerLeuThrLeuValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPhe 458

Qy       1678 CTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGG 1737
              |||      |||      :::   ||||||||||||||||||||   ||||||:::|||
Db        459 LeuTyrHisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArg 478

Qy       1738 GAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTA 1797
              |||:::||||||         ||||||   ||||||   |||:::||||||||||||:::|||
Db        479 GluGluAlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaVal 498
```

```
Qy       1798 GCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTGATCTTCCTTCT 1857
Db        499 AlaAlaArgGlyLeuAspIleSerAsnValArgHisValIleAsnPheAspLeuProSer 518

Qy       1858 ACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGA 1917
Db        519 AspIleGluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeu 538

Qy       1918 GCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTA 1977
Db        539 AlaThrSerPhePheAsn---GluLysAsnMetAsnIleThrLysAspLeuLeuAspLeu 557

Qy       1978 TTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATAC 2037
Db        558 LeuValGluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis 577

Qy       2038 ATTCCTGGCTTCAGT---GGTAGTACAAGAGGAAACGTGTTTGCA-----------TCA 2082
Db        578 TyrLysGlyGlySerArgGlyArgSerLysSerAsnArgPheSerGlyGlyPheGlyAla 597

Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
Db        598 ArgAspTyrArgGlnSerSerGlySer---SerSerSerGlyPheGlyAlaSerArg 615

RESULT 4
US-09-058-489-91
; Sequence 91, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 91
;   LENGTH: 660
;   TYPE: PRT
;   ORGANISM: Human
US-09-058-489-91

Alignment Scores:
Pred. No.:               6.01e-115      Length:        660
Score:                   1182.50        Matches:       278
Percent Similarity:      58.06%         Conservative:  93
Best Local Similarity:   43.51%         Mismatches:    167
Query Match:             29.91%         Indels:        101
DB:                      3              Gaps:          18

US-09-714-865-15 (1-2172) x US-09-058-489-91 (1-660)

Qy        301 TTTGAAGATGGTGATAGCTCTGGTTTCTGGAGAGAGTCTAGTAATGACTGC--------- 351
Db         52 PheHisAspLysAspSerSerGlyTrpSerCysSerLysAspLysAspAlaTyrSerSer 71

Qy        352 ---------GAAGATAATCCAACACGGAACAGAGGGTTTTCCAAGAGAGGCGGCTATCGA 402
Db         72 PheGlySerArgAspSerArgGlyLysProGlyTyrPheSerGluArgGlySer------ 89

Qy        403 GATGGAAATAATTCAGAAGCTTCAGGGCCATACAGAAGAGGTGGAAGAGGTAGTTTC--- 459
Db         90 ---------------------------------------GlySerArgGlyArgPheAsp 96

Qy        460 ---CGAGGTTGCCGTGGAGGATTTGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGAC 516
Db         97 AspArgGly-----------------------------ArgSerAspTyrAsp------ 104

Qy        517 GAATGTATGCAGCGCACTGGTGGCCTTTTTGGTTCTAGAAGACCAGTATTAAGTGGCACA 576
Db        105 ---------------------------------------------------GlyIle 106

Qy        577 GGTAATGGTGATACTTCTCAAAGCAGAAGTGGCAGTGGAAGT---GAACGAGGTGGTTAC 633
Db        107 GlyAsn-----------ArgGluArgProGlyPheGlyArgPheGluArgSerGlyHis 122

Qy        634 AAAGGTTTAAATGAAGAAGTAATAACAGGCTCTGGAAAGAATTCTTGGAAGTCAGAAGCA 693
Db        123 SerArgTrpCysAspLys-----------SerValGluAspAspTrpSerLys------ 136

Qy        694 GAAGGAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCA 753
Db        137 ---------------------------------------------ProLeuProPro 140

Qy        754 CCTGAGGAT---GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAA 810
Db        141 SerGluArgLeuGluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLys 160

Qy        811 TACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTT 870
Db        161 TyrAspAspIleProValGluAlaThrGlySerAsnCysProProHisIleGluAsnPhe 180

Qy        871 GAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAG 930
Db        181 SerAspIleAspMetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArg 200

Qy        931 CTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCT 990
Db        201 ProThrProValGlnLysHisAlaIleProIleIleLysGlyLysArgAspLeuValAla 220

Qy        991 TGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG 1050
Db        221 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIle 240

Qy       1051 ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA-------------- 1095
Db        241 TyrThrAspGly---ProGlyGluAlaLeuLysAlaValLysGluAsnGlyArgTyrGly 259

Qy       1096 ---------------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAG 1140
Db        260 ArgArgLysGlnTyrProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGln 279

Qy       1141 ATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATAT 1200
Db        280 IleTyrGluGluAlaArgLysPheSerTyrArgSerArgValArgProCysValValTyr 299

Qy       1201 GGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGT 1260
Db        300 GlyGlyAlaAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuVal 319

Qy       1261 GCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATC 1320
Db        320 AlaThrProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCys 339

Qy       1321 AAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATG 1380
Db        340 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIle 359

Qy       1381 AAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTC 1440
Db        360 ArgArgIleValGluGlnAspThrMetProProLysGlyValArgHisThrMetMetPhe 379

Qy       1441 AGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTAT 1500
Db        380 SerAlaThrPheProLysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyr 398

Qy       1501 CTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTC 1560
Db        399 IlePheLeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnLysValVal 418
```

```
Qy       1561 CAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG--- 1617
                 |||           ||||||      |||:::::::||||||      |||
Db        419 TrpValGluAspLeuAspLysArgSerPheLeuLeuAspIleLeuGlyAlaThrGlySer 438

Qy       1618 GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTT 1677
              |||         |||:::|||||||||||||||||||||      ||||||         |||
Db        439 AspSerLeuThrLeuValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPhe 458

Qy       1678 CTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGG 1737
              |||         |||        :::   ||||||||||||||||||||   ||||||:::|||
Db        459 LeuTyrHisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArg 478

Qy       1738 GAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTA 1797
              |||:::||||||      ||||||   ||||||   |||:::||||||||||||:::|||
Db        479 GluGluAlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaVal 498

Qy       1798 GCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCT 1857
              ||||||||||||||||||   ||||||:::||||||||||||||||||||||||||||||
Db        499 AlaAlaArgGlyLeuAspIleSerAsnValArgHisValIleAsnPheAspLeuProSer 518

Qy       1858 ACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGA 1917
                 |||:::||||||||||||||||||||||||||||||||   ||||||   |||
Db        519 AspIleGluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeu 538

Qy       1918 GCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTA 1977
              |||   |||||||||:::   |||   :::   ::::::   :::   |||:::   :::
Db        539 AlaThrSerPhePheAsn---GluLysAsnMetAsnIleThrLysAspLeuLeuAspLeu 557

Qy       1978 TTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTACATAC 2037
              |||   :::|||:::|||:::||||||:::|||||||||   :::|||:::      :::
Db        558 LeuValGluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaTyrGluHisHis 577

Qy       2038 ATTCCTGGCTTCAGT---GGTAGTACAAGAGGAAACGTGTTTGCA-----------TCA 2082
                 |||   |||   |||   ::::::   |||   |||:::               :::
Db        578 TyrLysGlyGlySerArgGlyArgSerLysSerAsnArgPheSerGlyGlyPheGlyAla 597

Qy       2083 GTTGATACCAGAAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGA 2139
                 |||   |||:::         :::   :::::::::||||||   :::||||||
Db        598 ArgAspTyrArgGlnSerSerGlySer---SerSerSerGlyPheGlyAlaSerArg 615
```

RESULT 5
US-09-183-706-43
; Sequence 43, Application US/09183706
; Patent No. 6245525
; GENERAL INFORMATION:
; APPLICANT: Martelange, Valrie
; APPLICANT: De Smet, Charles
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: L0461/7054
; CURRENT APPLICATION NUMBER: US/09/183,706
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 09/122,989
; EARLIER FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 43
; LENGTH: 648
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-183-706-43

Alignment Scores:

Pred. No.:	2.35e-65	Length:	648
Score:	709.50	Matches:	160
Percent Similarity:	60.00%	Conservative:	77
Best Local Similarity:	40.51%	Mismatches:	145
Query Match:	17.94%	Indels:	13
DB:	4	Gaps:	8

US-09-714-865-15 (1-2172) x US-09-183-706-43 (1-648)

```
Qy        850 CCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGT---CAGACACTGAATAACAAC 906
                 |||                ||||||:::::|||      |||            :::         |||
Db        237 ProAsnProThrCysThrPheAspAspAlaPheGlnCysTyrProGluValMetGluAsn 256

Qy        907 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 966
              |||   |||||||||:::   |||   ||||||:::|||         :::   ||||||:::
Db        257 IleLysLysAlaGlyPheGlnLysProThrProIleGlnSerGlnAlaTrpProIleVal 276

Qy        967 CTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||   |||   ||||||:::         ||||||||||||:::|||||||||      :::
Db        277 LeuGlnGlyIleAspLeuIleGlyValAlaGlnThrGlyThrGlyLysThrLeuCysTyr 296

Qy       1027 CTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG 1086
              |||:::|||      |||::::::                 :::         |||      :::
Db        297 LeuMetProGlyPheIleHisLeuVal---------LeuGlnProSerLeuLysGlyGln 313

Qy       1087 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
                 |||      :::::::::   ||||||||||||||||      |||:::
Db        314 ArgAsnArgProGlyMetLeuValLeuThrProThrArgGluLeuAlaLeuGlnValGlu 333

Qy       1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
                 |||      |||:::|||:::               :::|||:::|||      :::||||||||
Db        334 GlyGluCysCysLysTyrSerTyr---LysGlyLeuArgSerValCysValTyrGlyGly 352

Qy       1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
                          |||   ::::::   :::|||   :::|||:::   ||||||
Db        353 GlyAsnArgAspGluGlnIleGluGluLeuLysLysGlyValAspIleIleIleAlaThr 372

Qy       1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
              ||||||||||||   |||:::               :::   ||||||   |||   |||
Db        373 ProGlyArgLeuAsnAspLeuGlnMetSerAsnPheValAsnLeuLysAsnIleThrTyr 392

Qy       1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
              |||||||||||||||||||||:::||||||||||||||||||   |||::::::   |||
Db        393 LeuValLeuAspGluAlaAspLysMetLeuAspMetGlyPheGluProGlnIleMetLys 412

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
              ::::::              :::      :::      |||||||||:::|||   ||||||
Db        413 IleLeu------------LeuAspValArgProAspArgGlnThrValMetThrSerAla 428

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              |||:::|||      :::   |||||||||      :::||||||            :::
Db        429 ThrTrpProHisSerValHisArgLeuAlaGlnSerTyrLeuLysGluPro---MetIle 447

Qy       1507 GTTGCTGTTGGACAAGTGGGT---GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAA 1563
              |||   ||||||      :::         |||            |||:::|||   ::::::
Db        448 ValTyrValGlyThrLeuAspLeuValAlaValSerSerValLysGlnAsnIleIleVal 467

Qy       1564 GTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT--- 1620
                 :::      |||         :::         |||:::::::::
Db        468 ThrThrGluGluGluLysTrpSerHisMetGlnThrPheLeuGlnSerMetSerSerThr 487

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
              ::::::   :::|||||||||      |||      ||||||      ::::::   |||
Db        488 AspLysValIleValPheValSerArgLysAlaValAlaAspHisLeuSerSerAspLeu 507

Qy       1681 TGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
                          ||||||      |||:::||||||||||||||||||||||:::||||||
Db        508 IleLeuGlyAsnIleSerValGluSerLeuHisGlyAspArgGluGlnArgAspArgGlu 527

Qy       1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
              :::||||||   :::|||:::   ||||||      :::|||:::||||||   :::|||
Db        528 LysAlaLeuGluAsnPheLysThrGlyLysValArgIleLeuIleAlaThrAspLeuAla 547

Qy       1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
              :::||||||||||||:::   :::|||   ||||||   |||||||||   |||
Db        548 SerArgGlyLeuAspValHisAspValThrHisValTyrAsnPheAspPheProArgAsn 567

Qy       1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
              |||:::||||||||||||||||||||||||||||   |||   ||||||   :::
Db        568 IleGluGluTyrValHisArgIleGlyArgThrGlyArgAlaGlyArgThrGlyValSer 587

Qy       1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
              |||:::               :::|||      :::|||      |||:::   :::|||
Db        588 IleThrThrLeu---ThrArgAsnAspTrpArgValAlaSerGluLeuIleAsnIleLeu 606
```

```
Qy       1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCC 2025
                   |||  |||  :::|||    |||     :::|||
Db        607 GluArgAlaAsnGlnSerIleProGluGluLeuValSerMetAla 621

RESULT 6
US-09-567-995-43
; Sequence 43, Application US/09567995
; Patent No. 6303756
; GENERAL INFORMATION:
;  APPLICANT:  Martelange, Val,rie
;  APPLICANT:  De Smet, Charles
;  APPLICANT:  Boon-Falleur, Thierry
;  TITLE OF INVENTION:  TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE:  L0461/7054
;  CURRENT APPLICATION NUMBER: US/09/567,995
;  CURRENT FILING DATE:  2000-05-10
;  PRIOR APPLICATION NUMBER:  09/183,706
;  PRIOR FILING DATE:  1998-10-30
;  NUMBER OF SEQ ID NOS:  43
; SEQ ID NO 43
;   LENGTH: 648
;   TYPE: PRT
;   ORGANISM: H. sapiens
US-09-567-995-43

Alignment Scores:
Pred. No.:              2.35e-65       Length:        648
Score:                  709.50         Matches:       160
Percent Similarity:     60.00%         Conservative:  77
Best Local Similarity:  40.51%         Mismatches:    145
Query Match:            17.94%         Indels:        13
DB:                     4              Gaps:          8

US-09-714-865-15 (1-2172) x US-09-567-995-43 (1-648)

Qy        850 CCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGT---CAGACACTGAATAACAAC 906
              |||          ||||||::::::|||    |||          :::      |||
Db        237 ProAsnProThrCysThrPheAspAspAlaPheGlnCysTyrProGluValMetGluAsn 256

Qy        907 ATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATA 966
              |||  |||||||||:::   |||   ||||||:::|||      :::   ||||||:::
Db        257 IleLysLysAlaGlyPheGlnLysProThrProIleGlnSerGlnAlaTrpProIleVal 276

Qy        967 CTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTT 1026
              |||   |||   ||||||:::      ||||||||||||:::|||||||||      :::
Db        277 LeuGlnGlyIleAspLeuIleGlyValAlaGlnThrGlyThrGlyLysThrLeuCysTyr 296

Qy       1027 CTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAG 1086
              |||:::|||         |||::::::         :::      |||          :::
Db        297 LeuMetProGlyPheIleHisLeuVal---------LeuGlnProSerLeuLysGlyGln 313

Qy       1087 TTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTAT 1146
                    |||      :::::::::   |||||||||||||||      |||:::
Db        314 ArgAsnArgProGlyMetLeuValLeuThrProThrArgGluLeuAlaLeuGlnValGlu 333

Qy       1147 TTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGA 1206
                 |||      |||:::|||:::         :::|||:::|||    :::|||||||||
Db        334 GlyGluCysCysLysTyrSerTyr---LysGlyLeuArgSerValCysValTyrGlyGly 352

Qy       1207 ACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACT 1266
                                   |||   ::::::   :::|||   :::|||:::   ||||||
Db        353 GlyAsnArgAspGluGlnIleGluGluLeuLysLysGlyValAspIleIleIleAlaThr 372

Qy       1267 CCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATAC 1326
              ||||||||||||    |||:::            :::    ||||||   |||   |||
Db        373 ProGlyArgLeuAsnAspLeuGlnMetSerAsnPheValAsnLeuLysAsnIleThrTyr 392

Qy       1327 TTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAG 1386
              |||||||||||||||||||||:::||||||||||||||||||      |||::::::   |||
Db        393 LeuValLeuAspGluAlaAspLysMetLeuAspMetGlyPheGluProGlnIleMetLys 412

Qy       1387 TTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCA 1446
                  ::::::                 :::      :::      |||||||||:::|||   ||||||
Db        413 IleLeu------------LeuAspValArgProAspArgGlnThrValMetThrSerAla 428

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              |||:::|||         :::   |||||||||         :::||||||         :::
Db        429 ThrTrpProHisSerValHisArgLeuAlaGlnSerTyrLeuLysGluPro---MetIle 447

Qy       1507 GTTGCTGTTGGACAAGTGGGT---GGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAA 1563
              |||   ||||||    :::          |||         |||:::|||   ::::::
Db        448 ValTyrValGlyThrLeuAspLeuValAlaValSerSerValLysGlnAsnIleIleVal 467

Qy       1564 GTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT--- 1620
                 :::      |||         :::         |||:::::::::
Db        468 ThrThrGluGluGluLysTrpSerHisMetGlnThrPheLeuGlnSerMetSerSerThr 487

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
              ::::::   :::|||||||||      |||      ||||||      ::::::   |||
Db        488 AspLysValIleValPheValSerArgLysAlaValAlaAspHisLeuSerSerAspLeu 507

Qy       1681 TGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
                       ||||||      |||:::||||||||||||||||||||||||:::||||||
Db        508 IleLeuGlyAsnIleSerValGluSerLeuHisGlyAspArgGluGlnArgAspArgGlu 527

Qy       1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
              :::||||||   :::|||:::      ||||||   :::|||:::||||||      :::|||
Db        528 LysAlaLeuGluAsnPheLysThrGlyLysValArgIleLeuIleAlaThrAspLeuAla 547

Qy       1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
              :::||||||||||||:::   :::|||   ||||||   |||||||||   |||
Db        548 SerArgGlyLeuAspValHisAspValThrHisValTyrAsnPheAspPheProArgAsn 567

Qy       1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
              |||:::||||||||||||||||||||||||||||||   |||   ||||||   :::
Db        568 IleGluGluTyrValHisArgIleGlyArgThrGlyArgAlaGlyArgThrGlyValSer 587

Qy       1921 ATTTCCTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
              |||:::            :::|||      :::|||      |||:::   :::|||
Db        588 IleThrThrLeu---ThrArgAsnAspTrpArgValAlaSerGluLeuIleAsnIleLeu 606

Qy       1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCC 2025
                   |||  |||  :::|||    |||     :::|||
Db        607 GluArgAlaAsnGlnSerIleProGluGluLeuValSerMetAla 621

RESULT 7
US-09-208-742-4
; Sequence 4, Application US/09208742
; Patent No. 6174679
; GENERAL INFORMATION:
;  APPLICANT: Kaufmann, Joerg
;  TITLE OF INVENTION: CIF150/hTAFII150 is Necessary for Cell
;  TITLE OF INVENTION: Cycle Progression
;  FILE REFERENCE: 1453.002
;  CURRENT APPLICATION NUMBER: US/09/208,742
;  CURRENT FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 1261
;   TYPE: PRT
;   ORGANISM: human
US-09-208-742-4

Alignment Scores:
Pred. No.:              1.39e-64       Length:        1261
Score:                  704.00         Matches:       227
Percent Similarity:     41.98%         Conservative:  113
Best Local Similarity:  28.02%         Mismatches:    264
Query Match:            17.80%         Indels:        206
DB:                     4              Gaps:          26

US-09-714-865-15 (1-2172) x US-09-208-742-4 (1-1261)

Qy        139 GGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGA 198
```

```
Db     52 GlyProGlyThrLysArgGlyPheGlyPheGlyGlyPheAlaIle------SerAlaGly 69

Qy    199 AACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT---------------ACATCCACA 243

Db     70 LysLysGluGluProLysLeuProGlnGlnSerHisSerAlaPheGlyAlaThrSerSer 89

Qy    244 ATGGGTGGTTTTGGA-----------------------------------GTTGGAAAG 267

Db     90 SerSerGlyPheGlyLysSerAlaProProGlnLeuProSerPheTyrLysIleGlySer 109

Qy    268 AGTTTTGGAAACAGAGGTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTC 327

Db    110 LysArgAlaAsnPheAspGluGluAsnAlaTyrPheGluAspGluGlu----------- 125

Qy    328 TGGAGAGAGTCTAGTAAT------------------GACTGCGAAGATAATCCAACACGG 369

Db    126 ---GluAspSerSerAsnValAspLeuProTyrIleProAlaGluAsnSerProThrArg 144

Qy    370 AACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGG 429

Db    145 GlnGlnPheHisSerLys------------------------------------------ 150

Qy    430 CCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGA 489

Db    150 ------------------------------------------------------------ 150

Qy    490 AGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGT 549

Db    151 ProValAspSerAspSerAspAspAsp--------------------------------- 159

Qy    550 TCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGA------ 603

Db    160 ---------ProLeuGluAlaPheMetAlaGluValGluAspGlnAlaAlaArgAspMet 176

Qy    604 ---------AGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654

Db    177 LysArgLeuGluGluLysAspLysGluArgLysAsnValLysGlyIleArgAspAspIle 196

Qy    655 ATAACAGGCTCTGGAAAGAATTCTTGG--------------------------------- 681

Db    197 GluGluGluAspAspGlnGluAlaTyrPheArgTyrMetAlaGluAsnProThrAlaGly 216

Qy    682 ------AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGA------------ 723

Db    217 ValValGlnGluGluGluGluAspAsnLeuGluTyrAspSerAspGlyAsnProIleAla 236

Qy    724 CCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--------------------- 762

Db    237 ProThrLysLysIleIleAspProLeuProProIleAspHisSerGluIleAspTyrPro 256

Qy    763 ------GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAAC--------------- 801

Db    257 ProPheGluLysAsnPheTyrAsnGluHisGluGluIleThrAsnLeuThrProGlnGln 276

Qy    802 ---TTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCA 858

Db    277 LeuIleAspLeuArgHisLysLeuAsnLeuArgValSerGlyAlaAlaProProArgPro 296

Qy    859 ATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCT 918

Db    297 GlySerSerPheAlaHisPheGlyPheAspGluGlnLeuMetHisGlnIleArgLysSer 316

Qy    919 GGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGA 978

Db    317 GluTyrThrGlnProThrProIleGlnCysGlnGlyValProValAlaLeuSerGlyArg 336

Qy    979 GATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038

Db    337 AspMetIleGlyIleAlaLysThrGlySerGlyLysThrAlaAlaPheIleTrpProMet 356

Qy   1039 TTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA--- 1095

Db    357 LeuIleHisIleMetAspGln--------------------LysGluLeuGluProGly 369

Qy   1096 ------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149

Db    370 AspGlyProIleAlaValIleValCysProThrArgGluLeuCysGlnGlnIleHisPro 389

Qy   1150 GAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACC 1209

Db    390 GluCysLysArgPheGlyLysAlaTyrAsnLeuArgSerValAlaValTyrGlyGlyGly 409

Qy   1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269

Db    410 SerMetTrpGluGlnAlaLysAlaLeuGlnGluGlyAlaGluIleValValCysThrPro 429

Qy   1270 GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA 1329

Db    430 GlyArgLeuIleAspHisValLysLysLysAlaThrAsnLeuGlnArgValSerTyrLeu 449

Qy   1330 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTA 1389

Db    450 ValPheAspGluAlaAspArgMetPheAspMetGlyPheGluTyrGlnValArgSerIle 469

Qy   1390 ATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACT 1449

Db    470 AlaSer-----------HisValArgProAspArgGlnThrLeuLeuPheSerAlaThr 485

Qy   1450 TTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTT 1509

Db    486 PheArgLysLysIleGluLysLeuAlaArgAspIleLeu---IleAspProIleArgVal 504

Qy   1510 GCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT------CTCCAA 1563

Db    505 ValGlnGlyAspIleGlyGluAlaAsnGluAspValThrGlnIleValGluIleLeuHis 524

Qy   1564 GTTGGCCAGTTCTCAAAA------------AGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611

Db    525 SerGlyPro---SerLysTrpAsnTrpLeuThrArgArgLeuValGluPhe--------- 540

Qy   1612 ATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCA 1671

Db    541 ThrSerSerGlySerValLeuLeuPheValThrLysLysAlaAsnAlaGluGluLeuAla 560

Qy   1672 ACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGA 1731

Db    561 AsnAsnLeuLysGlnGluGlyHisAsnLeuGlyLeuLeuHisGlyAspMetAspGlnSer 580

Qy   1732 GAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACT 1791

Db    581 GluArgAsnLysValIleSerAspPheLysLysLysAspIleProValLeuValAlaThr 600

Qy   1792 TCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTT 1851

Db    601 AspValAlaAlaArgGlyLeuAspIleProSerIleLysThrValIleAsnTyrAspVal 620

Qy   1852 CCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACT 1911

Db    621 AlaArgAspIleAspThrHisThrHisArgIleGlyArgThrGlyArgAlaGlyGluLys 640

Qy   1912 GGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTA 1971

Db    641 GlyValAlaTyrThrLeuLeu---ThrProLysAspSerAsnPheAlaGlyAspLeuVal 659

Qy   1972 AAAGTATTGACAGATGCTCAACAGGATGTTCCT--------------------------- 2004

Db    660 ArgAsnLeuGluGlyAlaAsnGlnHisValSerLysGluLeuLeuAspLeuAlaMetGln 679

Qy   2005 ---GCATGGTTGGAAGAAATTGCCTTTAGT------------------------------ 2031

Db    680 AsnAlaTrpPheArgLysSerArgPheLysGlyGlyLysGlyLysLysLeuAsnIleGly 699

Qy   2032 -----------------ACATACATTCCTGGCTTCAGTGGTAGTACA------AGAGGA 2067

Db    700 GlyGlyGlyLeuGlyTyrArgGluArgProGlyLeuGlySerGluAsnMetAspArgGly 719
```

```
Qy     2068 AAC----------------------------------------------------------- 2070
Db      720 AsnAsnAsnValMetSerAsnTyrGluAlaTyrLysProSerThrGlyAlaMetGlyAsp 739

Qy     2071 ----------------------------------------------------GTGTTTGCA 2079
Db      740 ArgLeuThrAlaMetLysAlaAlaPheGlnSerGlnTyrLysSerHisPheValAlaAla 759

Qy     2080 TCAGTTGATACCAGAAAG---GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCA 2136
Db      760 SerLeuSerAsnGlnLysAlaGlySerSerAlaAlaGlyAlaSerGlyTrpThrSerAla 779

Qy     2137 CGAGCTCCCAATCCAGTAGATGATGAGTCA 2166
Db      780 GlySerLeuAsnSerValProThrAsnSer 789

RESULT 8
US-09-332-295-2
; Sequence 2, Application US/09332295
; Patent No. 6303372
; GENERAL INFORMATION:
;  APPLICANT: Kaufmann, Joerg
;  TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
;  FILE REFERENCE: 200130.456 / 1513.003
;  CURRENT APPLICATION NUMBER: US/09/332,295
;  CURRENT FILING DATE: 1999-06-11
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1261
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-332-295-2

Alignment Scores:
Pred. No.:              1.39e-64     Length:         1261
Score:                  704.00       Matches:        227
Percent Similarity:     41.98%       Conservative:   113
Best Local Similarity:  28.02%       Mismatches:     264
Query Match:            17.80%       Indels:         206
DB:                     4            Gaps:           26

US-09-714-865-15 (1-2172) x US-09-332-295-2 (1-1261)

Qy      139 GGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGA 198
Db       52 GlyProGlyThrLysArgGlyPheGlyPheGlyGlyPheAlaIle------SerAlaGly 69

Qy      199 AACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT---------------ACATCCACA 243
Db       70 LysLysGluGluProLysLeuProGlnGlnSerHisSerAlaPheGlyAlaThrSerSer 89

Qy      244 ATGGGTGGTTTTGGA-----------------------------------GTTGGAAAG 267
Db       90 SerSerGlyPheGlyLysSerAlaProProGlnLeuProSerPheTyrLysIleGlySer 109

Qy      268 AGTTTTGGAAACAGAGGTTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTC 327
Db      110 LysArgAlaAsnPheAspGluGluAsnAlaTyrPheGluAspGluGlu------------ 125

Qy      328 TGGAGAGAGTCTAGTAAT------------------GACTGCGAAGATAATCCAACACGG 369
Db      126 ---GluAspSerSerAsnValAspLeuProTyrIleProAlaGluAsnSerProThrArg 144

Qy      370 AACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGG 429
Db      145 GlnGlnPheHisSerLys------------------------------------------ 150

Qy      430 CCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGA 489
Db      150 ------------------------------------------------------------ 150

Qy      490 AGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGT 549
Db      151 ProValAspSerAspSerAspAspAsp--------------------------------- 159

Qy      550 TCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGA------ 603
Db      160 ---------ProLeuGluAlaPheMetAlaGluValGluAspGlnAlaAlaArgAspMet 176

Qy      604 ---------AGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654
Db      177 LysArgLeuGluGluLysAspLysGluArgLysAsnValLysGlyIleArgAspAspIle 196

Qy      655 ATAACAGGCTCTGGAAAGAATTCTTGG--------------------------------- 681
Db      197 GluGluGluAspAspGlnGluAlaTyrPheArgTyrMetAlaGluAsnProThrAlaGly 216

Qy      682 ------AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGA------------ 723
Db      217 ValValGlnGluGluGluGluAspAsnLeuGluTyrAspSerAspGlyAsnProIleAla 236

Qy      724 CCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--------------------- 762
Db      237 ProThrLysLysIleIleAspProLeuProProIleAspHisSerGluIleAspTyrPro 256

Qy      763 ------GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAAC--------------- 801
Db      257 ProPheGluLysAsnPheTyrAsnGluHisGluGluIleThrAsnLeuThrProGlnGln 276

Qy      802 ---TTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCA 858
Db      277 LeuIleAspLeuArgHisLysLeuAsnLeuArgValSerGlyAlaAlaProProArgPro 296

Qy      859 ATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCT 918
Db      297 GlySerSerPheAlaHisPheGlyPheAspGluGlnLeuMetHisGlnIleArgLysSer 316

Qy      919 GGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGA 978
Db      317 GluTyrThrGlnProThrProIleGlnCysGlnGlyValProValAlaLeuSerGlyArg 336

Qy      979 GATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
Db      337 AspMetIleGlyIleAlaLysThrGlySerGlyLysThrAlaAlaPheIleTrpProMet 356

Qy     1039 TTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA--- 1095
Db      357 LeuIleHisIleMetAspGln--------------------LysGluLeuGluProGly 369

Qy     1096 ------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149
Db      370 AspGlyProIleAlaValIleValCysProThrArgGluLeuCysGlnGlnIleHisPro 389

Qy     1150 GAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACC 1209
Db      390 GluCysLysArgPheGlyLysAlaTyrAsnLeuArgSerValAlaValTyrGlyGlyGly 409

Qy     1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269
Db      410 SerMetTrpGluGlnAlaLysAlaLeuGlnGluGlyAlaGluIleValValCysThrPro 429

Qy     1270 GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA 1329
Db      430 GlyArgLeuIleAspHisValLysLysLysAlaThrAsnLeuGlnArgValSerTyrLeu 449

Qy     1330 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTA 1389
Db      450 ValPheAspGluAlaAspArgMetPheAspMetGlyPheGluTyrGlnValArgSerIle 469

Qy     1390 ATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACT 1449
Db      470 AlaSer------------HisValArgProAspArgGlnThrLeuLeuPheSerAlaThr 485

Qy     1450 TTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTT 1509
```

```
                 |||    ::::::|||::::::|||||    :::    |||       :::   :::    |||
Db        486 PheArgLysLysIleGluLysLeuAlaArgAspIleLeu---IleAspProIleArgVal 504

Qy       1510 GCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT------CTCCAA 1563
                    |||   :::|||   |||     ||||||   |||   |||       |||
Db        505 ValGlnGlyAspIleGlyGluAlaAsnGluAspValThrGlnIleValGluIleLeuHis 524

Qy       1564 GTTGGCCAGTTCTCAAAA------------AGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611
                 |||      ||||||                      :::|||||||||
Db        525 SerGlyPro---SerLysTrpAsnTrpLeuThrArgArgLeuValGluPhe--------- 540

Qy       1612 ATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCA 1671
                          ::::::||||||         |||      |||:::        |||
Db        541 ThrSerSerGlySerValLeuLeuPheValThrLysLysAlaAsnAlaGluGluLeuAla 560

Qy       1672 ACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGA 1731
                    |||   ||||||       :::            :::||||||||    :::|||
Db        561 AsnAsnLeuLysGlnGluGlyHisAsnLeuGlyLeuLeuHisGlyAspMetAspGlnSer 580

Qy       1732 GAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACT 1791
              ||||||   :::   :::    ||||||:::             ||||||||||||||||||
Db        581 GluArgAsnLysValIleSerAspPheLysLysLysAspIleProValLeuValAlaThr 600

Qy       1792 TCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTT 1851
                 ||||||||||||||||||||||||     :::::::::    |||||||||:::|||:::
Db        601 AspValAlaAlaArgGlyLeuAspIleProSerIleLysThrValIleAsnTyrAspVal 620

Qy       1852 CCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACT 1911
                          ||||||   :::   ||||||||||||||||||||||||     |||
Db        621 AlaArgAspIleAspThrHisThrHisArgIleGlyArgThrGlyArgAlaGlyGluLys 640

Qy       1912 GGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTA 1971
              |||   |||   :::                    |||::::::   |||      ||||||
Db        641 GlyValAlaTyrThrLeuLeu---ThrProLysAspSerAsnPheAlaGlyAspLeuVal 659

Qy       1972 AAAGTATTGACAGATGCTCAACAGGATGTTCCT-------------------------- 2004
              :::    |||      |||    |||    |||
Db        660 ArgAsnLeuGluGlyAlaAsnGlnHisValSerLysGluLeuLeuAspLeuAlaMetGln 679

Qy       2005 ---GCATGGTTGGAAGAAATTGCCTTTAGT----------------------------- 2031
                 ||||||      :::      |||
Db        680 AsnAlaTrpPheArgLysSerArgPheLysGlyGlyLysGlyLysLysLeuAsnIleGly 699

Qy       2032 -----------------ACATACATTCCTGGCTTCAGTGGTAGTACA------AGAGGA 2067
                                         ||||||                   ||||||
Db        700 GlyGlyGlyLeuGlyTyrArgGluArgProGlyLeuGlySerGluAsnMetAspArgGly 719

Qy       2068 AAC-------------------------------------------------------- 2070
              |||
Db        720 AsnAsnAsnValMetSerAsnTyrGluAlaTyrLysProSerThrGlyAlaMetGlyAsp 739

Qy       2071 ---------------------------------------------------GTGTTTGCA 2079
                                                                 |||   |||
Db        740 ArgLeuThrAlaMetLysAlaAlaPheGlnSerGlnTyrLysSerHisPheValAlaAla 759

Qy       2080 TCAGTTGATACCAGAAAG---GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCA 2136
              |||:::      :::|||   |||   |||            :::|||::::::|||:::
Db        760 SerLeuSerAsnGlnLysAlaGlySerSerAlaAlaGlyAlaSerGlyTrpThrSerAla 779

Qy       2137 CGAGCTCCCAATCCAGTAGATGATGAGTCA 2166
                 :::   |||   |||         |||
Db        780 GlySerLeuAsnSerValProThrAsnSer 789
```

RESULT 9
US-09-709-979-2
; Sequence 2, Application US/09709979
; Patent No. 6423822
; GENERAL INFORMATION:
;   APPLICANT: Kaufmann, Joerg
;   TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
;   FILE REFERENCE: 200130.456 / 1513.003
;   CURRENT APPLICATION NUMBER: US/09/709,979
;   CURRENT FILING DATE: 2000-11-09
;   PRIOR APPLICATION NUMBER: US 09/332,295
;   PRIOR FILING DATE: 1999-06-11
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1261
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-709-979-2

```
Alignment Scores:
Pred. No.:                1.39e-64       Length:        1261
Score:                    704.00         Matches:       227
Percent Similarity:       41.98%         Conservative:  113
Best Local Similarity:    28.02%         Mismatches:    264
Query Match:              17.80%         Indels:        206
DB:                       4              Gaps:          26

US-09-714-865-15 (1-2172) x US-09-709-979-2 (1-1261)

Qy        139 GGACCTTCTCGAAGAGATCATTTCATGAAAAGTGGATTTGCCTCTGGGCGGAATTTTGGA 198
              ||||||          :::       |||         ||||||||||   :::   |||
Db         52 GlyProGlyThrLysArgGlyPheGlyPheGlyGlyPheAlaIle------SerAlaGly 69

Qy        199 AACAGAGATGCTGGTGAGTGTAATAAGCGAGATAAT---------------ACATCCACA 243
                 ::::::      :::      ::::::   :::                ||||||:::
Db         70 LysLysGluGluProLysLeuProGlnGlnSerHisSerAlaPheGlyAlaThrSerSer 89

Qy        244 ATGGGTGGTTTTGGA-----------------------------------GTTGGAAAG 267
                    |||||||||                                    :::|||
Db         90 SerSerGlyPheGlyLysSerAlaProProGlnLeuProSerPheTyrLysIleGlySer 109

Qy        268 AGTTTTGGAAACAGAGGTTTTCAAACAGCAGGTTTGAAGATGGTGATAGCTCTGGTTTC 327
                       |||            |||:::   |||||||||     :::
Db        110 LysArgAlaAsnPheAspGluGluAsnAlaTyrPheGluAspGluGlu------------ 125

Qy        328 TGGAGAGAGTCTAGTAAT------------------GACTGCGAAGATAATCCAACACGG 369
                 :::|||||||||                        |||::::::|||||||||
Db        126 ---GluAspSerSerAsnValAspLeuProTyrIleProAlaGluAsnSerProThrArg 144

Qy        370 AACAGAGGGTTTTCCAAGAGAGGCGGCTATCGAGATGGAAATAATTCAGAAGCTTCAGGG 429
                 :::      ||||||
Db        145 GlnGlnPheHisSerLys------------------------------------------ 150

Qy        430 CCATACAGAAGAGGTGGAAGAGGTAGTTTCCGAGGTTGCCGTGGAGGATTTGGTCTAGGA 489

Db        150 ------------------------------------------------------------ 150

Qy        490 AGTCCAAATAATGACTTAGACCCAGACGAATGTATGCAGCGCACTGGTGGCCTTTTTGGT 549
                    ::::::|||   |||   |||
Db        151 ProValAspSerAspSerAspAspAsp--------------------------------- 159

Qy        550 TCTAGAAGACCAGTATTAAGTGGCACAGGTAATGGTGATACTTCTCAAAGCAGA------ 603
                         |||:::   :::                 :::            :::|||
Db        160 ---------ProLeuGluAlaPheMetAlaGluValGluAspGlnAlaAlaArgAspMet 176

Qy        604 ---------AGTGGCAGTGGAAGTGAACGAGGTGGTTACAAAGGTTTAAATGAAGAAGTA 654
                                ||||||         ||||||:::   :::::::::
Db        177 LysArgLeuGluGluLysAspLysGluArgLysAsnValLysGlyIleArgAspAspIle 196

Qy        655 ATAACAGGCTCTGGAAAGAATTCTTGG--------------------------------- 681
                          :::   ::::::
Db        197 GluGluGluAspAspGlnGluAlaTyrPheArgTyrMetAlaGluAsnProThrAlaGly 216

Qy        682 ------AAGTCAGAAGCAGAAGGAGGAGAAAGTAGTGATACTCAAGGA------------ 723
                    :::   |||   |||                  |||:::   |||
Db        217 ValValGlnGluGluGluGluAspAsnLeuGluTyrAspSerAspGlyAsnProIleAla 236

Qy        724 CCAAAAGTGACCTACATACCCCCTCCTCCACCTGAGGAT--------------------- 762
                 |||         |||   |||   ||||||   |||
Db        237 ProThrLysLysIleIleAspProLeuProProIleAspHisSerGluIleAspTyrPro 256

Qy        763 ------GAGGACTCCATCTTTGCACATTATCAGACAGGCATAAAC--------------- 801
```

```
                 |||   :::   :::        ::::::        |||
Db        257 ProPheGluLysAsnPheTyrAsnGluHisGluGluIleThrAsnLeuThrProGlnGln 276

Qy        802 ---TTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCA 858
                 |||         :::   :::  ||||||||          |||
Db        277 LeuIleAspLeuArgHisLysLeuAsnLeuArgValSerGlyAlaAlaProProArgPro 296

Qy        859 ATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCT 918
                 :::|||          :::   |||   :::   |||   |||:::
Db        297 GlySerSerPheAlaHisPheGlyPheAspGluGlnLeuMetHisGlnIleArgLysSer 316

Qy        919 GGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGA 978
                 ||||||:::   |||||||:::|||          :::|||:::   |||:::||||||
Db        317 GluTyrThrGlnProThrProIleGlnCysGlnGlyValProValAlaLeuSerGlyArg 336

Qy        979 GATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATT 1038
              |||::::::      |||:::|||||||||||||||||||||||||||:::   |||:::
Db        337 AspMetIleGlyIleAlaLysThrGlySerGlyLysThrAlaAlaPheIleTrpProMet 356

Qy       1039 TTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAA--- 1095
              |||   |||:::|||                         |||||||||:::
Db        357 LeuIleHisIleMetAspGln---------------------LysGluLeuGluProGly 369

Qy       1096 ------CCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149
                    |||      :::||||||   |||||||||||||||      ||||||:::
Db        370 AspGlyProIleAlaValIleValCysProThrArgGluLeuCysGlnGlnIleHisPro 389

Qy       1150 GAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACC 1209
              |||   ::::::|||             :::|||:::|||   :::|||||||||
Db        390 GluCysLysArgPheGlyLysAlaTyrAsnLeuArgSerValAlaValTyrGlyGlyGly 409

Qy       1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269
                 :::         :::   :::   :::|||      |||:::      ||||||
Db        410 SerMetTrpGluGlnAlaLysAlaLeuGlnGluGlyAlaGluIleValValCysThrPro 429

Qy       1270 GGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTA 1329
              |||||||||:::|||   :::   |||:::         |||:::::::::   ||||||
Db        430 GlyArgLeuIleAspHisValLysLysLysAlaThrAsnLeuGlnArgValSerTyrLeu 449

Qy       1330 GTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTA 1389
              |||   ||||||||||||||||||   ||||||||||||      ::::::::   :::
Db        450 ValPheAspGluAlaAspArgMetPheAspMetGlyPheGluTyrGlnValArgSerIle 469

Qy       1390 ATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACT 1449
                 |||               :::   |||||||||||:::||||||||||||
Db        470 AlaSer------------HisValArgProAspArgGlnThrLeuLeuPheSerAlaThr 485

Qy       1450 TTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTT 1509
              |||   ::::::|||::::::||||||   :::   |||         :::   :::   |||
Db        486 PheArgLysLysIleGluLysLeuAlaArgAspIleLeu---IleAspProIleArgVal 504

Qy       1510 GCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT------CTCCAA 1563
                 |||   :::|||   |||      ||||||   |||   |||      |||
Db        505 ValGlnGlyAspIleGlyGluAlaAsnGluAspValThrGlnIleValGluIleLeuHis 524

Qy       1564 GTTGGCCAGTTCTCAAAA------------AGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611
                 |||      ||||||                 :::|||||||||
Db        525 SerGlyPro---SerLysTrpAsnTrpLeuThrArgArgLeuValGluPhe--------- 540

Qy       1612 ATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCA 1671
                          ::::::||||||      |||      |||:::         |||
Db        541 ThrSerSerGlySerValLeuLeuPheValThrLysLysAlaAsnAlaGluGluLeuAla 560

Qy       1672 ACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGA 1731
                 |||   ||||||      :::         :::|||||||||   :::|||
Db        561 AsnAsnLeuLysGlnGluGlyHisAsnLeuGlyLeuLeuHisGlyAspMetAspGlnSer 580

Qy       1732 GAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACT 1791
              ||||||   :::   :::   ||||||:::               ||||||||||||||||||
Db        581 GluArgAsnLysValIleSerAspPheLysLysLysAspIleProValLeuValAlaThr 600

Qy       1792 TCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTT 1851
                 ||||||||||||||||||||||||   :::::::::   |||||||||:::|||:::
Db        601 AspValAlaAlaArgGlyLeuAspIleProSerIleLysThrValIleAsnTyrAspVal 620

Qy       1852 CCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACT 1911
                          ||||||   :::   |||||||||||||||||||||   |||
Db        621 AlaArgAspIleAspThrHisThrHisArgIleGlyArgThrGlyArgAlaGlyGluLys 640

Qy       1912 GGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTA 1971
              |||   |||   :::                  |||::::::   |||   ||||||
Db        641 GlyValAlaTyrThrLeuLeu---ThrProLysAspSerAsnPheAlaGlyAspLeuVal 659

Qy       1972 AAAGTATTGACAGATGCTCAACAGGATGTTCCT--------------------------- 2004
              :::   |||         |||   |||   |||
Db        660 ArgAsnLeuGluGlyAlaAsnGlnHisValSerLysGluLeuLeuAspLeuAlaMetGln 679

Qy       2005 ---GCATGGTTGGAAGAAATTGCCTTTAGT------------------------------ 2031
                 ||||||      :::   |||
Db        680 AsnAlaTrpPheArgLysSerArgPheLysGlyGlyLysGlyLysLysLeuAsnIleGly 699

Qy       2032 ------------------ACATACATTCCTGGCTTCAGTGGTAGTACA------AGAGGA 2067
                                   ||||||                        ||||||
Db        700 GlyGlyGlyLeuGlyTyrArgGluArgProGlyLeuGlySerGluAsnMetAspArgGly 719

Qy       2068 AAC--------------------------------------------------------- 2070
              |||
Db        720 AsnAsnAsnValMetSerAsnTyrGluAlaTyrLysProSerThrGlyAlaMetGlyAsp 739

Qy       2071 ---------------------------------------------------GTGTTTGCA 2079
                                                                 |||   |||
Db        740 ArgLeuThrAlaMetLysAlaAlaPheGlnSerGlnTyrLysSerHisPheValAlaAla 759

Qy       2080 TCAGTTGATACCAGAAAG---GGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCA 2136
              |||:::      :::|||   |||   |||            :::|||::::::|||:::
Db        760 SerLeuSerAsnGlnLysAlaGlySerSerAlaAlaGlyAlaSerGlyTrpThrSerAla 779

Qy       2137 CGAGCTCCCAATCCAGTAGATGATGAGTCA 2166
                 :::   |||   |||         |||
Db        780 GlySerLeuAsnSerValProThrAsnSer 789
```

RESULT 10
US-09-134-001C-4262
; Sequence 4262, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4262
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4262

Alignment Scores:

Pred. No.:	2.03e-50	Length:	528
Score:	566.50	Matches:	136
Percent Similarity:	53.23%	Conservative:	78
Best Local Similarity:	33.83%	Mismatches:	153
Query Match:	14.33%	Indels:	35
DB:	4	Gaps:	8

US-09-714-865-15 (1-2172) x US-09-134-001C-4262 (1-528)

```
Qy        868 TTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACT 927
              |||:::|||         :::                  :::         |||:::
Db         23 PheLysGluLeuGlyIleSerAspLysThrValGlnThrLeuGluAlaMetGlyPheLys 42
```

```
Qy        928 AAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATG 987
              :::    |||||:::|||||    ||||||||       |||   .|||    |||:::::
Db         43 GluProThrProIleGlnLysAspSerIleProTyrAlaLeuGluGlyAspAspIleLeu 62

Qy        988 GCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTGGCTCAT 1047
                 |||||||||||:::||||||||   ||||||   :::|||::::::
Db         63 GlyGlnAlaGlnThrGlyThrGlyLysThrGlyAlaPheGlyIleProLeuIleGluLys 82

Qy       1048 ATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
              ::::::                                        |||     :::   :::
Db         83 ValValGlyGln-----------------------------GlnGlyValGlnSerLeu 92

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCT 1167
              |||:::|||||||||||||||||       |||:::     :::   |||:::||||||
Db         93 IleLeuAlaProThrArgGluLeuAlaMetGlnValAlaGluGlnLeuArgGluPheSer 112

Qy       1168 TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATT 1227
                 |||     |||:::   |||   ::::::||||||     :::           |||
Db        113 LysGlyGlnLysValGlnValValThrValPheGlyGlyMetProIleGluArgGlnIle 132

Qy       1228 CGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATC 1287
              :::   :::   :::|||      |||:::     ||||||||||||:::::::|||
Db        133 LysAlaLeuLysArgGlyProGlnIleValValGlyThrProGlyArgValIleAspHis 152

Qy       1288 ATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGAT 1347
              :::   :::      :::      :::   |||      |||:::||||||||||||||
Db        153 LeuAsnArgArgThrLeuLysThrGlnGlyIleHisThrLeuIleLeuAspGluAlaAsp 172

Qy       1348 CGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATG 1407
                 |||::::::|||||||||       :::|||:::   ::::::             :::
Db        173 GluMetMetAsnMetGlyPheIleAspAspMetArgPheIleMetAsp------LysIle 190

Qy       1408 CCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAA 1467
              |||:::   ||||||||||||||||::::::||||||||||||   |||:::   ||||||
Db        191 ProAla---GluGlnArgGlnThrMetLeuPheSerAlaThrMetProLysAlaIleGln 209

Qy       1468 AGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTT---GGACAAGTG 1524
                 |||      :::|||:::|||:::         :::         ::::::
Db        210 GluLeuValGlnGlnPheMetLysAlaProLysIleIleLysThrMetAsnAsnGluMet 229

Qy       1525 GGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGA 1584
                       |||      |||:::         |||   :::      |||
Db        230 SerAspProGlnIleAspGluTyrTyrThrIle------ValLysGluLeuGluLysPhe 247

Qy       1585 GAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAA 1644
              :::               |||            |||      :::||||||
Db        248 AspThrPheThrAsnPheLeuAspValHisGlnProGluLeuAlaIleValPheGlyArg 267

Qy       1645 ACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACA 1704
              ||||||::::::   |||            :::   |||      :::
Db        268 ThrLysArgArgValAspGluLeuThrSerAlaLeuLeuSerLysGlyTyrLysAlaGlu 287

Qy       1705 AGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTT 1764
                 :::|||||||||      |||      :::|||   :::   |||      |||:::
Db        288 GlyLeuHisGlyAspIleThrGlnAlaLysArgLeuGluValLeuLysLysPheLysAsn 307

Qy       1765 GGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAAT 1824
                 :::      :::||||||||||||   ||||||||||||||||||||||
Db        308 AspGlnIleAspIleLeuValAlaThrAspValAlaAlaArgGlyLeuAspIleSerGly 327

Qy       1825 GTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATT 1884
              |||   ||||||   |||||||||:::|||         :::   |||   ||||||||
Db        328 ValSerHisValTyrAsnPheAspIleProGlnAspThrGluSerTyrThrHisArgIle 347

Qy       1885 GGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGAT--------- 1935
              ||||||||||||||   |||      |||   |||:::::|||      :::
Db        348 GlyArgThrGlyArgAlaGlyLysGluGlyIleAlaValThrPheValAsnProIleGlu 367

Qy       1936 ------------------CTTGAATCGGATAACCAT---------------TTAGCACAG 1962
                                :::|||   |||:::               |||
Db        368 MetAspTyrIleArgGlnIleGluAspValAsnAsnArgArgMetLysAlaLeuArgPro 387

Qy       1963 CCTCTAGTAAAAGTATTGACAGATGCTCAACAGGAT---------------GTTCCTGCA 2007
              |||      |||   :::      |||::::::|||               |||
Db        388 ProHisArgLysGluValLeuLysAlaArgGluAspAspIleLysAspArgValGlnAsn 407

Qy       2008 TGGTTG 2013
              |||:::
Db        408 TrpMet 409
```

RESULT 11
US-09-318-443-8
; Sequence 8, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318,443
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-318-443-8

Alignment Scores:

Pred. No.:	4.27e-44	Length:	411
Score:	505.50	Matches:	130
Percent Similarity:	51.38%	Conservative:	93
Best Local Similarity:	29.95%	Mismatches:	150
Query Match:	12.78%	Indels:	61
DB:	4	Gaps:	10

US-09-714-865-15 (1-2172) x US-09-318-443-8 (1-411)

```
Qy        757 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGAC 816
              ::::::||||||                                    |||   :::   |||:::      :::
Db         19 LysGluGluAsp------------------MetThrLysValGluPheGluThrSerGlu 32

Qy        817 ACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAA 876
                 :::                  |||   |||         ||||||:::
Db         33 GluVal---------------------AspValThrPro---------ThrPheAspThr 42

Qy        877 GCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACT 936
                 |||   :::   |||         |||         |||:::   |||      :::
Db         43 MetGlyLeuArgGluAspLeuLeuArgGlyIleTyrAlaTyrGlyPheGluLysProSer 62

Qy        937 CCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCT 996
                 :::|||:::   :::|||      |||:::   ||||||||:::::::|||      :::
Db         63 AlaIleGlnGlnArgAlaIleLysGlnIleIleLysGlyArgAspValIleAlaGlnSer 82

Qy        997 CAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCAT 1056
              |||:::|||:::||||||||||||   |||   :::   :::|||         :::
Db         83 GlnSerGlyThrGlyLysThrAlaThrPheSerIleSerValLeuGlnCysLeu------ 100

Qy       1057 GATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCA 1116
                                         :::::::::|||   :::   :::|||:::|||
Db        101 ---------------------AspIleGlnValArgGluThrGlnAlaLeuIleLeuAla 113

Qy       1117 CCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACT 1176
              ||||||||||||||||          ||||||
Db        114 ProThrArgGluLeuAlaValGlnIleGlnLysGlyLeuLeuAlaLeuGlyAspTyrMet 133

Qy       1177 TGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATA 1236
                 |||:::            |||||||||   :::|||         ||||||::::::
Db        134 AsnValGlnCysHisAlaCysIleGlyGlyThrAsnValGlyGluAspIleArgLysLeu 153

Qy       1237 GTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAA 1296
                 |||   :::::::::         ||||||||||||:::   |||:::|||   :::
Db        154 AspTyrGlyGlnHisValValAlaGlyThrProGlyArgValPheAspMetIleArgArg 173
```

```
Qy       1297 GAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTG 1356
                     :::      :::   ||||||   ||||||||||||||||||||    ||||||
Db        174 ArgSerLeuArgThrArgAlaIleLysMetLeuValLeuAspGluAlaAspGluMetLeu 193

Qy       1357 GATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAG 1416
              :::   ||||||                                            |||
Db        194 AsnLysGlyPhe--------------------------------------------Lys 198

Qy       1417 GAACAG--------------------------------CGCCAAACCCTTATGTTCAGT 1443
              ||||||                                  |||  ::::::   |||
Db        199 GluGlnIleTyrAspValTyrArgTyrLeuProProAlaThrGlnValValLeuIleSer 218

Qy       1444 GCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTG 1503
              ||||||   |||   ||||||         :::     :::|||:::   ::::::   :::
Db        219 AlaThrLeuProHisGluIleLeuGluMetThrAsnLysPheMet---ThrAspProIle 237

Qy       1504 TTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAA 1563
                 :::   |||   :::                  ::::::|||      :::
Db        238 ArgIleLeuValLysArgAspGluLeuThrLeuGluGlyIleLysGlnPhePheValAla 257

Qy       1564 GTTGGCCAGTTCTCA---AAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGAT 1620
              |||   :::          |||   :::   |||   ::::::          :::
Db        258 ValGluArgGluGluTrpLysPheAspThrLeuCysAspLeuTyrAspThrLeuThrIle 277

Qy       1621 GAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
                 :::   ::::::|||      ||||||:::|||   |||:::             :::
Db        278 ThrGlnAlaValIlePheCysAsnThrLysArgLysValAspTrpLeuThrGluLysMet 297

Qy       1681 TGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAG 1740
                 :::         :::   :::|||:::|||||||||      |||:::|||||||||
Db        298 ArgGluAlaAsnPheThrValSerSerMetHisGlyAspMetProGlnLysGluArgGlu 317

Qy       1741 CAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCT 1800
                 :::      :::||||||   |||         ||||||::::::|||   |||
Db        318 SerIleMetLysGluPheArgSerGlyAlaSerArgValLeuIleSerThrAspValTrp 337

Qy       1801 GCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACC 1860
              ||||||||||||||:::      |||         :::||||||:::|||||||||:::
Db        338 AlaArgGlyLeuAspValProGlnValSerLeuIleIleAsnTyrAspLeuProAsnAsn 357

Qy       1861 ATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCA 1920
                 :::   |||:::|||||||||||||||:::||||||   |||      |||   |||
Db        358 ArgGluLeuTyrIleHisArgIleGlyArgSerGlyArgTyrGlyArgLysGlyValAla 377

Qy       1921 ATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTG 1980
              |||:::|||         :::::::::|||:::                  :::::::::|||
Db        378 IleAsnPhe------ValLysAsnAspAsp------------------IleArgIleLeu 389

Qy       1981 ACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATT 2022
                 |||   :::|||            ::::::|||:::
Db        390 ArgAspIleGluGlnTyrTyrSerThrGlnIleAspGluMet 403
```

RESULT 12
US-09-318-443-2
; Sequence 2, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318,443
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-318-443-2

```
Alignment Scores:
Pred. No.:                5.48e-44     Length:         415
Score:                    504.50       Matches:        131
Percent Similarity:       50.81%       Conservative:   89
Best Local Similarity:    30.25%       Mismatches:     152
Query Match:              12.76%       Indels:         61
DB:                       4            Gaps:           10

US-09-714-865-15 (1-2172) x US-09-318-443-2 (1-415)

Qy        760 GATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACT 819
              :::||||||         :::||||||
Db         23 GluGluAspMetThrThrValGluPheGlnThrSer------------------------ 34

Qy        820 ATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCT 879
                    ||||||         |||         |||         ||||||:::
Db         35 ------GluGluVal---------AspValThrPro---------ThrPheAspThrMet 46

Qy        880 AATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCT 939
                 |||   :::   |||         |||         |||:::   |||   :::
Db         47 GlyLeuArgGluAspLeuLeuArgGlyIleTyrAlaTyrGlyPheGluLysProSerAla 66

Qy        940 GTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAA 999
              :::|||:::   :::|||      |||:::   |||||||||::::::|||   :::|||
Db         67 IleGlnGlnLysAlaIleLysGlnIleIleLysGlyArgAspValIleAlaGlnSerGln 86

Qy       1000 ACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTGGCTCATATGATGCATGAT 1059
              :::|||:::||||||||||||   |||   :::   :::|||         :::
Db         87 SerGlyThrGlyLysThrAlaThrPheCysValSerValLeuGlnCysLeu--------- 103

Qy       1060 GGAATAACTGCCAGTCGTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCA 1119
                                 :::::::::|||   :::   :::|||:::||||||
Db        104 ------------------AspIleGlnIleArgGluThrGlnAlaLeuIleLeuAlaPro 117

Qy       1120 ACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGT 1179
              |||:::||||||      ||||||
Db        118 ThrLysGluLeuAlaArgGlnIleGlnLysValLeuLeuAlaLeuGlyAspTyrMetAsn 137

Qy       1180 GTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTA 1239
              |||:::                  |||||||||   :::|||      ||||||:::::
Db        138 ValGlnCysHisAlaCysIleGlyGlyThrAsnValGlyGluAspIleArgLysLeuAsp 157

Qy       1240 CAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAA 1299
                 |||   :::::::::      ||||||||||||:::   |||:::|||   :::
Db        158 TyrGlyGlnHisValValAlaGlyThrProGlyArgValPheAspMetIleArgArgArg 177

Qy       1300 AAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGAT 1359
                 :::      :::   ||||||   ||||||||||||||||||||   ||||||:::
Db        178 SerLeuArgThrArgAlaIleLysMetLeuValLeuAspGluAlaAspGluMetLeuAsn 197

Qy       1360 ATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAA 1419
                 ||||||                                          ||||||
Db        198 LysGlyPhe---------------------------------------------LysGlu 202

Qy       1420 CAG--------------------------------CGCCAAACCCTTATGTTCAGTGCA 1446
              |||                                 |||   :::   ||||||
Db        203 GlnIleTyrAspValTyrArgTyrLeuProProAlaThrGlnValCysLeuIleSerAla 222

Qy       1447 ACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTT 1506
              |||   |||   ||||||         :::      :::|||:::   ::::::   :::
Db        223 ThrLeuProHisGluIleLeuGluMetThrAsnLysPheMet---ThrAspProIleArg 241

Qy       1507 GTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTT 1566
              :::   |||   :::                  ::::::|||      :::   |||
Db        242 IleLeuValLysArgAspGluLeuThrLeuGluGlyIleLysGlnPhePheValAlaVal 261

Qy       1567 GGCCAGTTCTCA---AAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAA 1623
                 :::         |||   :::   |||   ::::::          :::
Db        262 GluArgGluGluTrpLysPheAspThrLeuCysAspLeuTyrAspThrLeuThrIleThr 281

Qy       1624 AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGT 1683
              :::   ::::::|||      ||||||:::|||   |||:::             :::
Db        282 GlnAlaValIlePheCysAsnThrLysArgLysValAspTrpLeuThrGluLysMetArg 301
```

```
Qy       1684 CAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAA 1743
               :::       :::   :::|||:::|||||||||        |||:::|||||||||
Db        302 GluAlaAsnPheThrValSerSerMetHisGlyAspMetProGlnLysGluArgGluSer 321

Qy       1744 GCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCC 1803
                 :::   :::||||||   |||         ||||||::::::|||   |||   |||
Db        322 IleMetLysGluPheArgSerGlyAlaSerArgValLeuIleSerThrAspValTrpAla 341

Qy       1804 AGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATT 1863
              ||||||||||||:::         |||      :::||||||:::|||||||||:::
Db        342 ArgGlyLeuAspValProGlnValSerLeuIleIleAsnTyrAspLeuProAsnAsnArg 361

Qy       1864 GATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATT 1923
              :::   |||:::|||||||||||||||:::||||||   |||      |||   ||||||
Db        362 GluLeuTyrIleHisArgIleGlyArgSerGlyArgTyrGlyArgLysGlyValAlaIle 381

Qy       1924 TCCTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACA 1983
              :::|||      :::::::::|||:::                   :::::::::|||
Db        382 AsnPhe------ValLysAsnAspAsp-----------------IleArgIleLeuArg 393

Qy       1984 GATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATT 2022
              |||   :::|||               ::::::|||:::
Db        394 AspIleGluGlnTyrTyrSerThrGlnIleAspGluMet 406
```

RESULT 13
US-09-318-443-6
; Sequence 6, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318,443
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-318-443-6

Alignment Scores:

Pred. No.:	6.93e-44	Length:	411
Score:	503.50	Matches:	128
Percent Similarity:	53.43%	Conservative:	98
Best Local Similarity:	30.26%	Mismatches:	158
Query Match:	12.73%	Indels:	39
DB:	4	Gaps:	10

US-09-714-865-15 (1-2172) x US-09-318-443-6 (1-411)

```
Qy        757 GAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGAC 816
              ::::::||||||                    |||   :::   |||:::      :::
Db         19 LysGluGluAsp------------------MetThrLysValGluPheGluThrSerGlu 32

Qy        817 ACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAA 876
              :::                       |||   |||            ||||||:::
Db         33 GluVal---------------------AspValThrPro---------ThrPheAspThr 42

Qy        877 GCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACT 936
                       |||   :::   |||         |||         |||:::   |||   :::
Db         43 MetGlyLeuArgGluAspLeuLeuArgGlyIleTyrAlaTyrGlyPheGluLysProSer 62

Qy        937 CCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCT 996
                 :::|||:::   :::|||      |||:::   |||||||||::::::|||   :::
Db         63 AlaIleGlnGlnArgAlaIleLysGlnIleIleLysGlyArgAspValIleAlaGlnSer 82

Qy        997 CAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCAT 1056
              |||:::|||:::|||||||||||      |||   :::   :::|||         :::
Db         83 GlnSerGlyThrGlyLysThrAlaThrPheSerIleSerValLeuGlnCysLeu------ 100

Qy       1057 GATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCA 1116
                                    :::::::::|||   :::   :::|||:::|||
Db        101 ---------------------AspIleGlnValArgGluThrGlnAlaLeuIleLeuAla 113

Qy       1117 CCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACT 1176
              ||||||||||||||||||         ||||||
Db        114 ProThrArgGluLeuAlaValGlnIleGlnLysGlyLeuLeuAlaLeuGlyAspTyrMet 133

Qy       1177 TGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATA 1236
                 |||:::               |||||||||   :::|||      ||||||::::::
Db        134 AsnValGlnCysHisAlaCysIleGlyGlyThrAsnValGlyGluAspIleArgLysLeu 153

Qy       1237 GTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAA 1296
                 |||   :::::::::      ||||||||||||:::   |||:::|||      :::
Db        154 AspTyrGlyGlnHisValValAlaGlyThrProGlyArgValPheAspMetIleArgArg 173

Qy       1297 GAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTG 1356
                 :::      :::   ||||||   |||||||||||||||||||||   ||||||
Db        174 ArgSerLeuArgThrArgAlaIleLysMetLeuValLeuAspGluAlaAspGluMetLeu 193

Qy       1357 GATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAG 1416
              :::   ||||||      ::::::      :::               :::||||||
Db        194 AsnLysGlyPheLysGluGlnIleTyrAspValTyrArgTyr------LeuProSer--- 210

Qy       1417 GAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCT 1476
                    |||   ::::::   |||||||||   |||   ||||||         :::
Db        211 ---AlaThrGlnValValLeuIleSerAlaThrLeuProHisGluIleLeuGluMetThr 229

Qy       1477 GCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGT 1536
                 :::||||||:::   ::::::   :::   :::   |||   :::
Db        230 AsnLysPheMet---ThrAspProIleArgIleLeuValLysArgAspGluLeuThrLeu 248

Qy       1537 AGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCA---AAAAGAGAAAAGCTC 1593
                 ::::::|||   :::   |||   :::            |||   :::   |||
Db        249 GluGlyIleLysGlnPhePheValAlaValGluArgGluGluTrpLysPheAspThrLeu 268

Qy       1594 GTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAA 1653
                 ::::::      :::         :::   ::::::|||      ||||||:::
Db        269 CysAspLeuTyrAspThrLeuThrIleThrGlnAlaValIlePheCysAsnThrLysArg 288

Qy       1654 AAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCAT 1713
                 |||   |||:::         :::   :::         :::   :::|||:::|||
Db        289 LysValAspTrpLeuThrGluLysMetArgGluAlaAsnPheThrValSerSerMetHis 308

Qy       1714 GGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGC 1773
              ||||||      |||:::|||||||||   :::   :::||||||   |||
Db        309 GlyAspMetProGlnLysGluArgGluSerIleMetLysGluPheArgSerGlyAlaSer 328

Qy       1774 CCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACAT 1833
                 ||||||::::::|||   |||   |||||||||||||||:::         |||
Db        329 ArgValLeuIleSerThrAspValTrpAlaArgGlyLeuAspValProGlnValSerLeu 348

Qy       1834 GTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACT 1893
              :::||||||:::|||||||||:::         :::   |||:::||||||||||||:::
Db        349 IleIleAsnTyrAspLeuProAsnAsnArgGluLeuTyrIleHisArgIleGlyArgSer 368

Qy       1894 GGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCAT 1953
              |||:::   |||      |||   ||||||:::|||         :::::::::|||:::
Db        369 GlyGlnTyrGlyArgLysGlyValAlaIleAsnPhe------ValLysAsnAspAsp--- 385

Qy       1954 TTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTG 2013
                             :::::::::|||   |||   :::|||               :::
Db        386 ---------------IleArgIleLeuArgAspIleGluGlnTyrTyrSerThrGlnIle 400

Qy       2014 GAAGAAATT 2022
              :::|||:::
Db        401 AspGluMet 403
```

RESULT 14
US-08-607-509-4
; Sequence 4, Application US/08607509

```
; Patent No. 5876735
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    TITLE OF INVENTION:  METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/607,509
;      FILING DATE:  16-FEB-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.404C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206)682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-607-509-4

Alignment Scores:
Pred. No.:               2.89e-42      Length:         403
Score:                   488.00        Matches:        126
Percent Similarity:      50.92%        Conservative:   96
Best Local Similarity:   28.90%        Mismatches:     164
Query Match:             12.34%        Indels:         50
DB:                      2             Gaps:           11

US-09-714-865-15 (1-2172) x US-08-607-509-4 (1-403)

Qy        754 CCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATAC 813
              |||:::|||:::||||||   |||         |||  |||:::
Db          9 ProGlnAspGlnAspSer---PheLeuAspAspGlnProGlyValArg----------- 23

Qy        814 GACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAA 873
                                                        |||   :::|||:::
Db         24 -----------------------------------------ProIleProSerPheAsp 29

Qy        874 GAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTT 933
              :::      |||   |||   |||         |||      |||:::   |||
Db         30 AspMetProLeuHisGlnAsnLeuLeuArgGlyIleTyrSerTyrGlyPheGluLysPro 49

Qy        934 ACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGT 993
              :::   :::|||:::   :::|||               |||   |||::::::|||
Db         50 SerSerIleGlnGlnArgAlaIleAlaProPheThrArgGlyGlyAspIleIleAlaGln 69

Qy        994 GCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATG--- 1050
              ||||||:::|||:::|||||||||   ||||||   :::   :::|||      :::
Db         70 AlaGlnSerGlyThrGlyLysThrGlyAlaPheSerIleGlyLeuLeuGlnArgLeuAsp 89

Qy       1051 ---ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
                 |||:::   |||                               :::   :::
Db         90 PheArgHisAsnLeuIle-------------------------------GlnGlyLeu 98

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAG---------ATTTATTTGGAAGCCAGA 1158
              :::::::::||||||||||||||     |||            :::
Db         99 ValLeuSerProThrArgGluLeuAlaLeuGlnThrAlaGluValIleSerArgIleGly 118

Qy       1159 AAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGGGA 1218
              :::|||            :::                     |||||||||::::::
Db        119 GluPheLeuSerAsnSerSerLysPheCysGluThrPheValGlyGlyThrArgValGln 138

Qy       1219 CATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTG 1278
                 :::|||::::::      |||      :::         ||||||||||||:::
Db        139 AspAspLeuArgLysLeuGlnAlaGlyValIleValAlaValGlyThrProGlyArgVal 158

Qy       1279 ATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTGGAT 1338
              |||:::|||   :::      :::      :::   ::::::   ||||||||||||
Db        159 SerAspValIleLysArgGlyAlaLeuArgThrGluSerLeuArgValLeuValLeuAsp 178

Qy       1339 GAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGC 1398
              |||||||||   ||||||      ||||||      ::::::   ::::::
Db        179 GluAlaAspGluMetLeuSerGlnGlyPheAlaAspGlnIleTyrGluIlePheArgPhe 198

Qy       1399 CCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAG 1458
                 :::|||      :::      |||      :::||||||||||||   ||||||
Db        199 ------LeuPro------LysAspIleGlnValAlaLeuPheSerAlaThrMetProGlu 214

Qy       1459 GAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGA 1518
              |||:::      |||      :::|||::::::   :::   :::   :::   |||
Db        215 GluValLeuGluLeuThrLysLysPheMetArg---AspProValArgIleLeuValLys 233

Qy       1519 CAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCA 1578
              :::                  ::::::|||      :::   |||   :::
Db        234 ArgGluSerLeuThrLeuGluGlyIleLysGlnPhePheIleAlaValGluGluGluHis 253

Qy       1579 AAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTT 1638
              |||   :::   |||:::::::::            :::         ::::::::::::|||
Db        254 LysLeuAspThrLeuMetAspLeuTyrGluThrValSerIleAlaGlnSerValIlePhe 273

Qy       1639 GTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCA 1698
                 |||::::::|||   |||:::      |||   |||         :::
Db        274 AlaAsnThrArgArgLysValAspTrpIleAlaGluLysLeuAsnGlnSerAsnHisThr 293

Qy       1699 ACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTT 1758
                 :::|||:::|||   :::      :::   :::||||||:::   :::      |||
Db        294 ValSerSerMetHisAlaGluMetProLysSerAspArgGluArgValMetAsnThrPhe 313

Qy       1759 CGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATT 1818
              |||   |||         |||||||||   |||   :::   |||||||||:::|||:::
Db        314 ArgSerGlySerSerArgValLeuValThrThrAspLeuValAlaArgGlyIleAspVal 333

Qy       1819 GAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCAT 1878
                 :::|||      ||||||||||||||||||||:::         :::   |||:::|||
Db        334 HisHisValAsnIleValIleAsnPheAspLeuProThrAsnLysGluAsnTyrLeuHis 353

Qy       1879 CGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTT 1938
              ||||||||||||   ||||||      |||      ||||||:::|||
Db        354 ArgIleGlyArgGlyGlyArgTyrGlyArgLysGlyValAlaIleAsnPheVal---Thr 372

Qy       1939 GAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGAT 1998
              |||   |||                              |||::::::|||   :::   :::
Db        373 GluLysAsp--------------------ValGluLeuLeuHisGluIleGluAlaHis 385

Qy       1999 GTTCCTGCATGGTTGGAAGAA------ATTGCCTTTAGTACATACATT 2040
                             ::::::|||      :::   |||:::   |||:::
Db        386 TyrHisThrGlnIleAspGluLeuProValAspPheAlaAlaTyrLeu 401

RESULT 15
US-08-634-642-4
; Sequence 4, Application US/08634642
; Patent No. 5879687
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    TITLE OF INVENTION:  METHODS FOR ENHANCEMENT OF
;    TITLE OF INVENTION:  PROTECTIVE IMMUNE RESPONSES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
```

```
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;.   COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,642
;      FILING DATE:  18-APR-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.404C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206)682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-634-642-4

Alignment Scores:
Pred. No.:               2.89e-42       Length:        403
Score:                   488.00         Matches:       126
Percent Similarity:      50.92%         Conservative:  96
Best Local Similarity:   28.90%         Mismatches:    164
Query Match:             12.34%         Indels:        50
DB:                      2              Gaps:          11

US-09-714-865-15 (1-2172) x US-08-634-642-4 (1-403)

Qy        754 CCTGAGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATAC 813
              |||:::|||:::|||||    |||        |||  |||:::
Db          9 ProGlnAspGlnAspSer---PheLeuAspAspGlnProGlyValArg----------- 23

Qy        814 GACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAA 873
                                                        |||   :::|||:::
Db         24 ------------------------------------------ProIleProSerPheAsp 29

Qy        874 GAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTT 933
              :::      |||   |||   |||          |||        |||:::   |||
Db         30 AspMetProLeuHisGlnAsnLeuLeuArgGlyIleTyrSerTyrGlyPheGluLysPro 49

Qy        934 ACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGT 993
              :::   :::|||:::   :::|||                |||   |||::::::|||
Db         50 SerSerIleGlnGlnArgAlaIleAlaProPheThrArgGlyGlyAspIleIleAlaGln 69

Qy        994 GCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTGGCTCATATG--- 1050
              ||||||:::|||:::||||||||   ||||||   :::   :::|||      :::
Db         70 AlaGlnSerGlyThrGlyLysThrGlyAlaPheSerIleGlyLeuLeuGlnArgLeuAsp 89

Qy       1051 ---ATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATT 1107
                 |||:::   |||                              :::   :::
Db         90 PheArgHisAsnLeuIle--------------------------------GlnGlyLeu 98

Qy       1108 ATTGTAGCACCAACTCGAGAATTGGTCAACCAG---------ATTTATTTGGAAGCCAGA 1158
              :::::::::|||||||||||||||      |||            :::
Db         99 ValLeuSerProThrArgGluLeuAlaLeuGlnThrAlaGluValIleSerArgIleGly 118

Qy       1159 AAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGA 1218
              :::|||          :::                     |||||||||::::::
Db        119 GluPheLeuSerAsnSerSerLysPheCysGluThrPheValGlyGlyThrArgValGln 138

Qy       1219 CATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTG 1278
                 :::|||::::::      |||       :::         ||||||||||||:::
Db        139 AspAspLeuArgLysLeuGlnAlaGlyValIleValAlaValGlyThrProGlyArgVal 158

Qy       1279 ATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGAT 1338
                 |||:::|||   :::      :::      :::   ::::::   ||||||||||||
Db        159 SerAspValIleLysArgGlyAlaLeuArgThrGluSerLeuArgValLeuValLeuAsp 178

Qy       1339 GAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGC 1398
              |||||||||   ||||||      ||||||      ::::::   ::::::
Db        179 GluAlaAspGluMetLeuSerGlnGlyPheAlaAspGlnIleTyrGluIlePheArgPhe 198

Qy       1399 CCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAG 1458
                    :::|||      :::      |||      :::||||||||||||   ||||||
Db        199 ------LeuPro------LysAspIleGlnValAlaLeuPheSerAlaThrMetProGlu 214

Qy       1459 GAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGA 1518
              |||:::      |||      :::|||::::::   :::   :::   :::   |||
Db        215 GluValLeuGluLeuThrLysLysPheMetArg---AspProValArgIleLeuValLys 233

Qy       1519 CAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCA 1578
              :::                  ::::::|||      :::   |||   :::
Db        234 ArgGluSerLeuThrLeuGluGlyIleLysGlnPhePheIleAlaValGluGluGluHis 253

Qy       1579 AAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTT 1638
              |||   :::   |||:::::::::            :::         ::::::::::::|||
Db        254 LysLeuAspThrLeuMetAspLeuTyrGluThrValSerIleAlaGlnSerValIlePhe 273

Qy       1639 GTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCA 1698
                    |||::::::|||   |||:::   |||         |||   |||      :::
Db        274 AlaAsnThrArgArgLysValAspTrpIleAlaGluLysLeuAsnGlnSerAsnHisThr 293

Qy       1699 ACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTT 1758
                 :::|||:::|||   :::         :::   :::||||||:::   :::   |||
Db        294 ValSerSerMetHisAlaGluMetProLysSerAspArgGluArgValMetAsnThrPhe 313

Qy       1759 CGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATT 1818
              |||   |||      |||||||||   |||   :::   |||||||||:::|||:::
Db        314 ArgSerGlySerSerArgValLeuValThrThrAspLeuValAlaArgGlyIleAspVal 333

Qy       1819 GAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCAT 1878
                 :::|||      |||||||||||||||||||||:::      :::   |||:::|||
Db        334 HisHisValAsnIleValIleAsnPheAspLeuProThrAsnLysGluAsnTyrLeuHis 353

Qy       1879 CGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTT 1938
              ||||||||||||   ||||||   |||         |||   ||||||:::|||
Db        354 ArgIleGlyArgGlyGlyArgTyrGlyArgLysGlyValAlaIleAsnPheVal---Thr 372

Qy       1939 GAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGAT 1998
              |||   |||                     |||::::::|||   :::   :::
Db        373 GluLysAsp---------------------ValGluLeuLeuHisGluIleGluAlaHis 385

Qy       1999 GTTCCTGCATGGTTGGAAGAA------ATTGCCTTTAGTACATACATT 2040
                       ::::::|||      :::   |||:::   |||:::
Db        386 TyrHisThrGlnIleAspGluLeuProValAspPheAlaAlaTyrLeu 401

Search completed: June 10, 2003, 16:45:29
Job time : 80 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:          June 10, 2003, 07:51:10 ; Search time 130 Seconds
                                           (without alignments)
                                           5123.861 Million cell updates/sec

Title:           US-09-714-865-15
Perfect score:   2172
Sequence:        1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	383.2	17.6	2319	3	US-09-058-489-90	Sequence 90, Appl
	2	383.2	17.6	4416	3	US-09-058-489-17	Sequence 17, Appl
	3	369.4	17.0	3408	3	US-09-058-489-14	Sequence 14, Appl
	4	369.4	17.0	5322	3	US-09-058-489-13	Sequence 13, Appl
	5	158.2	7.3	2365	4	US-09-183-706-42	Sequence 42, Appl
	6	158.2	7.3	2365	4	US-09-567-995-42	Sequence 42, Appl
	7	123.6	5.7	1254	5	PCT-US96-05320A-894	Sequence 894, App
c	8	116.8	5.4	3230	4	US-08-961-527-203	Sequence 203, App
	9	115.2	5.3	3825	4	US-09-208-742-3	Sequence 3, Appli
	10	103.2	4.8	1587	4	US-09-134-001C-1425	Sequence 1425, Ap
	11	92.2	4.2	1245	4	US-09-318-443-1	Sequence 1, Appli
	12	91.2	4.2	1368	4	US-09-134-001C-816	Sequence 816, App
	13	90.8	4.2	490	4	US-09-222-575-126	Sequence 126, App
	14	84.6	3.9	1347	2	US-08-959-749-1	Sequence 1, Appli
	15	84.6	3.9	1347	4	US-09-351-497-1	Sequence 1, Appli
	16	76.2	3.5	1374	3	US-08-929-738-1	Sequence 1, Appli
	17	76.2	3.5	1614	3	US-08-929-738-2	Sequence 2, Appli
c	18	73.6	3.4	7218	1	US-08-232-463-14	Sequence 14, Appl
	19	73	3.4	1800	3	US-09-039-773A-1	Sequence 1, Appli
	20	72.8	3.4	1682	4	US-09-318-443-7	Sequence 7, Appli
	21	71.2	3.3	1536	4	US-09-318-443-5	Sequence 5, Appli
	22	70.6	3.3	8494	4	US-08-961-527-163	Sequence 163, App
c	23	65.8	3.0	4774	4	US-09-221-017B-1	Sequence 1, Appli
	24	58.8	2.7	987	4	US-09-221-017B-1115	Sequence 1115, Ap
	25	58.2	2.7	2934	3	US-09-149-934-2	Sequence 2, Appli
	26	55.2	2.5	43676	3	US-09-356-952-12	Sequence 12, Appl
	27	55	2.5	2061	4	US-09-008-271A-16	Sequence 16, Appl
	28	55	2.5	2061	4	US-09-705-448-2	Sequence 2, Appli
	29	53.2	2.4	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	30	53.2	2.4	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	31	52.8	2.4	1867	2	US-08-607-509-3	Sequence 3, Appli
	32	52.8	2.4	1867	2	US-08-634-642-3	Sequence 3, Appli
	33	52.8	2.4	1867	3	US-08-989-370-3	Sequence 3, Appli
	34	50	2.3	299	4	US-09-118-554-42	Sequence 42, Appl
	35	50	2.3	299	4	US-09-118-627-42	Sequence 42, Appl
	36	50	2.3	299	4	US-09-602-877A-42	Sequence 42, Appl
	37	50	2.3	1380	4	US-09-134-001C-341	Sequence 341, App
	38	49.6	2.3	1618	2	US-08-533-669A-9	Sequence 9, Appli
	39	49.6	2.3	1618	2	US-08-607-509-1	Sequence 1, Appli
	40	49.6	2.3	1618	2	US-08-454-036-1	Sequence 1, Appli
	41	49.6	2.3	1618	2	US-08-634-642-1	Sequence 1, Appli
	42	49.6	2.3	1618	3	US-08-989-370-1	Sequence 1, Appli
	43	49.6	2.3	1618	4	US-09-183-861-9	Sequence 9, Appli
	44	49.6	2.3	1618	4	US-09-022-765-9	Sequence 9, Appli
	45	49.6	2.3	1618	5	PCT-US95-05064-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-058-489-90
; Sequence 90, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
;  APPLICANT: Whitehead Institute for Biomedical Research
;  APPLICANT: Lahn, Bruce
;  APPLICANT: Page, David
;  TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
;  TITLE OF INVENTION: the Y Chromosome
;  FILE REFERENCE: WHI97-08pA
;  CURRENT APPLICATION NUMBER: US/09/058,489
;  CURRENT FILING DATE: 1998-04-10
;  EARLIER APPLICATION NUMBER: 60/041,877
;  EARLIER FILING DATE: 1997-04-11
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 90
;   LENGTH: 2319
;   TYPE: DNA
;   ORGANISM: Human
US-09-058-489-90

  Query Match            17.6%;  Score 383.2;  DB 3;  Length 2319;
  Best Local Similarity  58.5%;  Pred. No. 4.6e-101;
  Matches  761;  Conservative    0;  Mismatches  503;  Indels   36;  Gaps    4;

Qy        788 AGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATG 847
               | || || || ||||| || ||||| ||    ||| |  || || |    | ||   | |
Db        529 ACACGGGGATTAACTTTGAGAAATATGATGATATACCAGTAGAGGCAACCGGCAGTAACT 588

Qy        848 CACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACA 907
                 || |||    |||  || ||||   ||   | |   |   |  | |   | |     ||||
Db        589 GTCCTCCACATATTGAGAATTTTAGCGATATTGACATGGGAGAAATTATCATGGGGAACA 648

Qy        908 TTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATAC 967
              |||        || | ||||||   | |||||| |||||||| |      |||||||| ||
Db        649 TTGAACTTACTCGCTATACTCGTCCTACTCCAGTGCAAAAACATGCCATTCCTATTATTA 708

Qy        968 TTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTC 1027
               |  |  | |||| ||   |||||||||| |||||||| ||||||||| ||||| || ||||
Db        709 AGGGAAAAAGAGACTTAGTGGCTTGTGCCCAAACAGGATCTGGGAAAACTGCAGCATTTC 768

Qy       1028 TCCTACCAATTTTGGCTCA--------------------------TATGATGCATGATG 1060
               |  ||||  ||  ||  |||                          | |||  |   ||
Db        769 TTTTACCCATACTGAGTCAGATATATACAGATGGTCCAGGAGAAGCTTTGAAGGCTGTGA 828

Qy       1061 GAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAA 1120
                  || ||  ||    |               ||  |  |||   |    |  || |||| ||||
Db        829 AGGAAAATGGAAGGTATGGGCGCCGCAAACAATATCCAATATCCTTGGTTTTAGCCCCAA 888
```

```
Qy       1121 CTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTG 1180
Db        889 CAAGAGAATTGGCTGTACAGATCTATGAGGAAGCCAGAAAATTTTCCTACCGATCTAGAG 948

Qy       1181 TAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTAC 1240
Db        949 TTCGTCCTTGTGTAGTTTATGGTGGTGCTGATATTGGTCAGCAGATTCGGGACTTAGAAC 1008

Qy       1241 AAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAA 1300
Db       1009 GTGGATGCCACTTGTTAGTAGCCACTCCAGGACGTCTAGTGGATATGATGGAAAGAGGAA 1068

Qy       1301 AGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATGTTGGATA 1360
Db       1069 AGATTGGATTAGACTTCTGCAAGTACTTAGTGTTGGATGAAGCTGATAGGATGCTGGATA 1128

Qy       1361 TGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAAC 1420
Db       1129 TGGGATTTGAACCTCAGATACGTCGTATAGTTGAACAAGATACTATGCCACCAAAGGGCG 1188

Qy       1421 AGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAG 1480
Db       1189 TTCGTCACACCATGATGTTTAGTGCTACTTTTCCTAAGGAAATACAGATGCTTGCTCGTG 1248

Qy       1481 AGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAG 1540
Db       1249 ACTTTTTGGATGAA---TATATCTTTTTGGCTGTAGGCAGAGTAGGCTCTACCTCTGAGA 1305

Qy       1541 ATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAA 1600
Db       1306 ACATCACACAGAAAGTAGTTTGGGTGGAAGACTTAGATAAACGGTCATTTCTACTGGACA 1365

Qy       1601 TTCTGCGAAACATAGGG---GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAG 1657
Db       1366 TTTTAGGTGCAACAGGGAGTGATTCACTTACTTTAGTGTTTGTGGAGACCAAAAAAGGGAG 1425

Qy       1658 CAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTG 1717
Db       1426 CAGATTCCCTGGAGGATTTCTTATACCATGAAGGATATGCTTGTACTAGTATTCATGGAG 1485

Qy       1718 ATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAG 1777
Db       1486 ACCGGTCACAGAGAGATCGAGAGGAGGCCCTTCACCAGTTTCGCTCAGGAAAAAGCCCAA 1545

Qy       1778 TTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTA 1837
Db       1546 TTCTAGTGGCTACAGCTGTGGCAGCACGAGGACTAGACATTTCAAATGTGAGACATGTTA 1605

Qy       1838 TCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTC 1897
Db       1606 TCAATTTTGATTTGCCAAGTGATATTGAAGAATATGTGCATCGTATTGGCCGTACAGGAC 1665

Qy       1898 GTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAG 1957
Db       1666 GTGTAGGAAACCTGGGCCTTGCCACCTCATTCTTTAA---TGAAAAAAATATGAATATTA 1722

Qy       1958 CACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAG 2017
Db       1723 CAAAGGATTTGTTGGATCTTCTTGTAGAAGCTAAACAAGAAGTGCCTTCTTGGTTGGAAA 1782

Qy       2018 AAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAG 2057
Db       1783 ATATGGCTTATGAACACCACTACAAGGGTGGCAGTCGTGG 1822
```

RESULT 2
US-09-058-489-17
; Sequence 17, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; TITLE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
;   LENGTH: 4416
;   TYPE: DNA
;   ORGANISM: Human
US-09-058-489-17

  Query Match   17.6%; Score 383.2; DB 3; Length 4416;
  Best Local Similarity   58.5%; Pred. No. 6.5e-101;
  Matches 761; Conservative   0; Mismatches 503; Indels   36; Gaps   4;

```
Qy        788 AGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATG 847
Db        529 ACACGGGGATTAACTTTGAGAAATATGATGATATACCAGTAGAGGCAACCGGCAGTAACT 588

Qy        848 CACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACA 907
Db        589 GTCCTCCACATATTGAGAATTTTAGCGATATTGACATGGGAGAAATTATCATGGGGAACA 648

Qy        908 TTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATAC 967
Db        649 TTGAACTTACTCGCTATACTCGTCCTACTCCAGTGCAAAAACATGCCATTCCTATTATTA 708

Qy        968 TTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTC 1027
Db        709 AGGGAAAAAGAGACTTAGTGGCTTGTGCCCAAACAGGATCTGGGAAAACTGCAGCATTTC 768

Qy       1028 TCCTACCAATTTTGGCTCA--------------------------TATGATGCATGATG 1060
Db        769 TTTTACCCATACTGAGTCAGATATATACAGATGGTCCAGGAGAAGCTTTGAAGGCTGTGA 828

Qy       1061 GAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAA 1120
Db        829 AGGAAAATGGAAGGTATGGGCGCCGCAAACAATATCCAATATCCTTGGTTTTAGCCCCAA 888

Qy       1121 CTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTG 1180
Db        889 CAAGAGAATTGGCTGTACAGATCTATGAGGAAGCCAGAAAATTTTCCTACCGATCTAGAG 948

Qy       1181 TAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTAC 1240
Db        949 TTCGTCCTTGTGTAGTTTATGGTGGTGCTGATATTGGTCAGCAGATTCGGGACTTAGAAC 1008

Qy       1241 AAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAA 1300
Db       1009 GTGGATGCCACTTGTTAGTAGCCACTCCAGGACGTCTAGTGGATATGATGGAAAGAGGAA 1068

Qy       1301 AGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATGTTGGATA 1360
Db       1069 AGATTGGATTAGACTTCTGCAAGTACTTAGTGTTGGATGAAGCTGATAGGATGCTGGATA 1128

Qy       1361 TGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAAC 1420
Db       1129 TGGGATTTGAACCTCAGATACGTCGTATAGTTGAACAAGATACTATGCCACCAAAGGGCG 1188

Qy       1421 AGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAG 1480
Db       1189 TTCGTCACACCATGATGTTTAGTGCTACTTTTCCTAAGGAAATACAGATGCTTGCTCGTG 1248

Qy       1481 AGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAG 1540
Db       1249 ACTTTTTGGATGAA---TATATCTTTTTGGCTGTAGGCAGAGTAGGCTCTACCTCTGAGA 1305

Qy       1541 ATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAA 1600
```

```
Db       1306 ACATCACACAGAAAGTAGTTTGGGTGGAAGACTTAGATAAACGGTCATTTCTACTGGACA 1365

Qy       1601 TTCTGCGAAACATAGGG---GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAG 1657

Db       1366 TTTTAGGTGCAACAGGGAGTGATTCACTTACTTTAGTGTTTGTGGAGACCAAAAAGGGAG 1425

Qy       1658 CAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTG 1717

Db       1426 CAGATTCCCTGGAGGATTTCTTATACCATGAAGGATATGCTTGTACTAGTATTCATGGAG 1485

Qy       1718 ATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAG 1777

Db       1486 ACCGGTCACAGAGAGATCGAGAGGAGGCCCTTCACCAGTTTCGCTCAGGAAAAAGCCCAA 1545

Qy       1778 TTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTA 1837

Db       1546 TTCTAGTGGCTACAGCTGTGGCAGCACGAGGACTAGACATTTCAAATGTGAGACATGTTA 1605

Qy       1838 TCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTC 1897

Db       1606 TCAATTTTGATTTGCCAAGTGATATTGAAGAATATGTGCATCGTATTGGCCGTACAGGAC 1665

Qy       1898 GTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAG 1957

Db       1666 GTGTAGGAAACCTGGGCCTTGCCACCTCATTCTTTAA---TGAAAAAAATATGAATATTA 1722

Qy       1958 CACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAG 2017

Db       1723 CAAAGGATTTGTTGGATCTTCTTGTAGAAGCTAAACAAGAAGTGCCTTCTTGGTTGGAAA 1782

Qy       2018 AAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAG 2057

Db       1783 ATATGGCTTATGAACACCACTACAAGGGTGGCAGTCGTGG 1822
```

RESULT 3
US-09-058-489-14
; Sequence 14, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; TITLE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
;   LENGTH: 3408
;   TYPE: DNA
;   ORGANISM: Human
US-09-058-489-14

```
  Query Match             17.0%;  Score 369.4;  DB 3;  Length 3408;
  Best Local Similarity   57.6%;  Pred. No. 5.7e-97;
  Matches  758;  Conservative    0;  Mismatches  521;  Indels   36;  Gaps    4;

Qy        773 TCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAG 832

Db       1305 TCTTTTCTGGAGGCAACACTGGGATTAATTTTGAGAAATACGATGACATTCCAGTTGAGG 1364

Qy        833 TGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGA 892

Db       1365 CAACAGGCAACAACTGTCCTCCACATATTGAAAGTTTCAGTGATGTTGAGATGGGAGAAA 1424

Qy        893 CACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACA 952

Db       1425 TTATCATGGGAAACATTGAGCTTACTCGTTATACTCGCCCAACTCCAGTGCAAAAGCATG 1484

Qy        953 GTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGA 1012

Db       1485 CTATTCCTATTATCAAAGAGAAAAGAGACTTGATGGCTTGTGCCCAAACAGGGTCTGGAA 1544

Qy       1013 AGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGG----------- 1061

Db       1545 AAACTGCAGCATTTCTGTTGCCCATCTTGAGTCAGATTTATTCAGATGGTCCAGGCGAGG 1604

Qy       1062 ----------------AATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTA 1105

Db       1605 CTTTGAGGGCCATGAAGGAAAATGGAAGGTATGGGCGCCGCAAACAATACCCAATCTCCT 1664

Qy       1106 TTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTT 1165

Db       1665 TGGTATTAGCACCAACGAGAGAGTTGGCAGTACAGATCTACGAAGAAGCCAGAAAATTTT 1724

Qy       1166 CTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAA 1225

Db       1725 CATACCGATCTAGAGTTCGTCCTTGCGTGGTTATGGTGGTGCCGATATTGGTCAGCAGA 1784

Qy       1226 TTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATA 1285

Db       1785 TTCGAGACTTGGAACGTGGATGCCATTTGTTAGTAGCCACTCCAGGACGTCTAGTGGATA 1844

Qy       1286 TCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTG 1345

Db       1845 TGATGGAAAGAGGAAAGATTGGATTAGACTTTTGCAAATACTTGGTGTTAGATGAAGCTG 1904

Qy       1346 ATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAA 1405

Db       1905 ATCGGATGTTGGATATGGGGTTTGAGCCTCAGATTCGTAGAATAGTCGAACAAGATACTA 1964

Qy       1406 TGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTC 1465

Db       1965 TGCCTCCAAAGGGTGTCCGCCACACTATGATGTTTAGTGCTACTTTTCCTAAGGAAATAC 2024

Qy       1466 AAAGGTTGGCTGCAGAGTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGG 1525

Db       2025 AGATGCTGGCTCGTGATTTCTTAGATGAA---TATATCTTCTTGGCTGTAGGAAGAGTTG 2081

Qy       1526 GTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAGAG 1585

Db       2082 GCTCTACCTCTGAAAACATCACACAGAAAGTAGTTTGGGTGGAAGAATCAGACAAACGGT 2141

Qy       1586 AAAAGCTCGTTGAAATTCTGCGAAACATAGG---GGATGAAAGAACTATGGTCTTTGTTG 1642

Db       2142 CATTTCTGCTTGACCTCCTAAATGCAACAGGCAAGGATTCACTGACCTTAGTGTTTGTGG 2201

Qy       1643 AAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTA 1702

Db       2202 AGACCAAAAAGGGTGCAGATTCTCTGGAGGATTTCTTATACCATGAAGGATACGCATGTA 2261

Qy       1703 CAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCT 1762

Db       2262 CCAGCATCCATGGAGACCGTTCTCAGAGGGATAGAGAAGAGGCCCTTCACCAGTTCCGCT 2321

Qy       1763 TTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAA 1822

Db       2322 CAGGAAAAAGCCCAATTTTAGTGGCTACAGCAGTAGCAGCAAGAGGACTGGACATTTCAA 2381

Qy       1823 ATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAA 1882

Db       2382 ATGTGAAACATGTTATCAATTTTGACTTGCCAAGTGATATTGAAGAATATGTACATCGTA 2441

Qy       1883 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAAT 1942

Db       2442 TTGGTCGTACGGGACGTGTAGGAAACCTTGGCCTGGCAACCTCATTCTTTAA---CGAGA 2498

Qy       1943 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 2002

Db       2499 GGAACATAAATATTACTAAGGATTTGTTGGATCTTCTTGTTGAAGCTAAACAAGAAGTGC 2558
```

```
Qy       2003 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAG 2057
Db       2559 CGTCTTGGTTAGAAAACATGGCTTATGAACACCACTACAAGGGTAGCAGTCGTGG 2613
```

RESULT 4
US-09-058-489-13
; Sequence 13, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; TITLE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WHI97-08pA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 5322
;   TYPE: DNA
;   ORGANISM: Human
US-09-058-489-13

```
  Query Match             17.0%;  Score 369.4;  DB 3;  Length 5322;
  Best Local Similarity   57.6%;  Pred. No. 7.2e-97;
  Matches  758;  Conservative    0;  Mismatches  521;  Indels   36;  Gaps    4;

Qy        773 TCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAG 832
Db       1305 TCTTTTCTGGAGGCAACACTGGGATTAATTTTGAGAAATACGATGACATTCCAGTTGAGG 1364

Qy        833 TGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGA 892
Db       1365 CAACAGGCAACAACTGTCCTCCACATATTGAAAGTTTCAGTGATGTTGAGATGGGAGAAA 1424

Qy        893 CACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACA 952
Db       1425 TTATCATGGGAAACATTGAGCTTACTCGTTATACTCGCCCAACTCCAGTGCAAAAGCATG 1484

Qy        953 GTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGA 1012
Db       1485 CTATTCCTATTATCAAAGAGAAAAGAGACTTGATGGCTTGTGCCCAAACAGGGTCTGGAA 1544

Qy       1013 AGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGG----------- 1061
Db       1545 AAACTGCAGCATTTCTGTTGCCCATCTTGAGTCAGATTTATTCAGATGGTCCAGGCGAGG 1604

Qy       1062 ----------------AATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTA 1105
Db       1605 CTTTGAGGGCCATGAAGGAAAATGGAAGGTATGGGCGCCGCAAACAATACCCAATCTCCT 1664

Qy       1106 TTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTT 1165
Db       1665 TGGTATTAGCACCAACGAGAGAGTTGGCAGTACAGATCTACGAAGAAGCCAGAAAATTTT 1724

Qy       1166 CTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAA 1225
Db       1725 CATACCGATCTAGAGTTCGTCCTTGCGTGGTTTATGGTGGTGCCGATATTGGTCAGCAGA 1784

Qy       1226 TTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATA 1285
Db       1785 TTCGAGACTTGGAACGTGGATGCCATTTGTTAGTAGCCACTCCAGGACGTCTAGTGGATA 1844

Qy       1286 TCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTG 1345
Db       1845 TGATGGAAAGAGGAAAGATTGGATTAGACTTTTGCAAATACTTGGTGTTAGATGAAGCTG 1904

Qy       1346 ATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAA 1405
Db       1905 ATCGGATGTTGGATATGGGGTTTGAGCCTCAGATTCGTAGAATAGTCGAACAAGATACTA 1964

Qy       1406 TGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTC 1465
Db       1965 TGCCTCCAAAGGGTGTCCGCCACACTATGATGTTTAGTGCTACTTTTCCTAAGGAAATAC 2024

Qy       1466 AAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGG 1525
Db       2025 AGATGCTGGCTCGTGATTTCTTAGATGAA---TATATCTTCTTGGCTGTAGGAAGAGTTG 2081

Qy       1526 GTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAG 1585
Db       2082 GCTCTACCTCTGAAAACATCACACAGAAAGTAGTTTGGGTGGAAGAATCAGACAAACGGT 2141

Qy       1586 AAAAGCTCGTTGAAATTCTGCGAAACATAGG---GGATGAAAGAACTATGGTCTTTGTTG 1642
Db       2142 CATTTCTGCTTGACCTCCTAAATGCAACAGGCAAGGATTCACTGACCTTAGTGTTTGTGG 2201

Qy       1643 AAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTA 1702
Db       2202 AGACCAAAAAGGGTGCAGATTCTCTGGAGGATTTCTTATACCATGAAGGATACGCATGTA 2261

Qy       1703 CAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCT 1762
Db       2262 CCAGCATCCATGGAGACCGTTCTCAGAGGGATAGAGAAGAGGCCCTTCACCAGTTCCGCT 2321

Qy       1763 TTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAA 1822
Db       2322 CAGGAAAAAGCCCAATTTTAGTGGCTACAGCAGTAGCAGCAAGAGGACTGGACATTTCAA 2381

Qy       1823 ATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAA 1882
Db       2382 ATGTGAAACATGTTATCAATTTTGACTTGCCAAGTGATATTGAAGAATATGTACATCGTA 2441

Qy       1883 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAAT 1942
Db       2442 TTGGTCGTACGGGACGTGTAGGAAACCTTGGCCTGGCAACCTCATTCTTTAA---CGAGA 2498

Qy       1943 CGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTC 2002
Db       2499 GGAACATAAATATTACTAAGGATTTGTTGGATCTTCTTGTTGAAGCTAAACAAGAAGTGC 2558

Qy       2003 CTGCATGGTTGGAAGAAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAG 2057
Db       2559 CGTCTTGGTTAGAAAACATGGCTTATGAACACCACTACAAGGGTAGCAGTCGTGG 2613
```

RESULT 5
US-09-183-706-42
; Sequence 42, Application US/09183706
; Patent No. 6245525
; GENERAL INFORMATION:
; APPLICANT: Martelange, Valrie
; APPLICANT: De Smet, Charles
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: L0461/7054
; CURRENT APPLICATION NUMBER: US/09/183,706
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 09/122,989
; EARLIER FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 42
;   LENGTH: 2365
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (208)...(2151)
US-09-183-706-42

```
  Query Match             7.3%;  Score 158.2;  DB 4;  Length 2365;
  Best Local Similarity   50.1%;  Pred. No. 8e-36;
  Matches  518;  Conservative    0;  Mismatches  488;  Indels   27;  Gaps    4;

Qy        902 ACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTA 961
Db        971 AAAACATTAAAAAGGCAGGTTTTCAAAAGCCAACACCTATTCAGTCACAGGCATGGCCCA 1030

Qy        962 TCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGG 1021
Db       1031 TTGTGTTGCAAGGAATAGATCTTATAGGAGTAGCCCAGACTGGAACAGGAAAGACATTGT 1090

Qy       1022 CTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTA 1081
Db       1091 GTTATTTAATGCCTGGATTTATTCATCTGGTCCTTCAAC---------CCAGCCTTAAAG 1141

Qy       1082 AAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGA 1141
Db       1142 GTCAAAGGAATAGACCCGGCATGTTAGTTCTAACTCCCACTCGGGAATTAGCACTTCAAG 1201

Qy       1142 TTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATG 1201
Db       1202 TAGAAGGAGAATGTTGCAAATATTCATAT---AAAGGGCTTCGGAGTGTTTGTGTATATG 1258

Qy       1202 GGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTG 1261
Db       1259 GTGGTGGAAATAGAGATGAACAAATAGAAGAGCTTAAAAAAGGTGTAGATATCATAATTG 1318

Qy       1262 CTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCA 1321
Db       1319 CAACTCCCGGAAGATTGAATGATCTGCAAATGAGTAACTTCGTCAATCTGAAGAATATAA 1378

Qy       1322 AATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGA 1381
Db       1379 CCTACTTGGTTTTAGATGAAGCAGACAAGATGTTGGACATGGGATTTGAACCCCAGATAA 1438

Qy       1382 AGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCA 1441
Db       1439 TGAAGATTTTGT------------TAGATGTGCGCCCAGATAGGCAGACAGTTATGACCA 1486

Qy       1442 GTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTT---AAAGTCAAATT 1498
Db       1487 GTGCTACATGGCCTCATTCAGTTCATCGCCTCGCACAATCTTATTTGAAAGAACCAATGA 1546

Qy       1499 ATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTC 1558
Db       1547 TTGTCTATGTTGGTACATTGGATCTAGTTGCTGTAAGTTCAGTGAAGCAAAATATAATTG 1606

Qy       1559 TCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGG 1618
Db       1607 TAACCACCGAGGAAGAGAAATGGAGTCACATGCAAACTTTTCTACAGAGTATGTCATCCA 1666

Qy       1619 ATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTC 1678
Db       1667 CAGACAAAGTCATTGTCTTCGTTTCTCGAAAAGCTGTTGCGGATCACTTATCAAGTGACC 1726

Qy       1679 TTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGG 1738
Db       1727 TAATACTTGGAAATATATCAGTAGAGTCTCTGCATGGAGATAGAGAACAGAGAGATCGGG 1786

Qy       1739 AGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAG 1798
Db       1787 AGAAAGCATTAGAGAACTTTAAAACAGGCAAAGTGAGAATACTAATTGCAACTGATCTAG 1846

Qy       1799 CTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTA 1858
Db       1847 CCTCTAGAGGACTTGATGTCCATGACGTTACACATGTCTATAATTTTGACTTTCCACGGA 1906

Qy       1859 CCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAG 1918
Db       1907 ATATTGAAGAATACGTACACCGAATAGGGCGCACGGGAAGAGCAGGGAGGACTGGTGTTT 1966

Qy       1919 CAATTTCCTTTTT 1931
Db       1967 CCATTACAACTTT 1979
```

RESULT 6

```
US-09-567-995-42
; Sequence 42, Application US/09567995
; Patent No. 6303756
; GENERAL INFORMATION:
;  APPLICANT:  Martelange, Val,rie
;  APPLICANT:  De Smet, Charles
;  APPLICANT:  Boon-Falleur, Thierry
;  TITLE OF INVENTION:  TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE:  L0461/7054
;  CURRENT APPLICATION NUMBER: US/09/567,995
;  CURRENT FILING DATE:  2000-05-10
;  PRIOR APPLICATION NUMBER:  09/183,706
;  PRIOR FILING DATE:  1998-10-30
;  NUMBER OF SEQ ID NOS:  43
; SEQ ID NO 42
;   LENGTH: 2365
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (208)...(2151)
US-09-567-995-42

  Query Match             7.3%;  Score 158.2;  DB 4;  Length 2365;
  Best Local Similarity   50.1%;  Pred. No. 8e-36;
  Matches  518;  Conservative    0;  Mismatches  488;  Indels   27;  Gaps    4;

Qy        902 ACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTA 961
Db        971 AAAACATTAAAAAGGCAGGTTTTCAAAAGCCAACACCTATTCAGTCACAGGCATGGCCCA 1030

Qy        962 TCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGG 1021
Db       1031 TTGTGTTGCAAGGAATAGATCTTATAGGAGTAGCCCAGACTGGAACAGGAAAGACATTGT 1090

Qy       1022 CTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTA 1081
Db       1091 GTTATTTAATGCCTGGATTTATTCATCTGGTCCTTCAAC---------CCAGCCTTAAAG 1141

Qy       1082 AAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGA 1141
Db       1142 GTCAAAGGAATAGACCCGGCATGTTAGTTCTAACTCCCACTCGGGAATTAGCACTTCAAG 1201

Qy       1142 TTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATG 1201
Db       1202 TAGAAGGAGAATGTTGCAAATATTCATAT---AAAGGGCTTCGGAGTGTTTGTGTATATG 1258

Qy       1202 GGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTG 1261
Db       1259 GTGGTGGAAATAGAGATGAACAAATAGAAGAGCTTAAAAAAGGTGTAGATATCATAATTG 1318

Qy       1262 CTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCA 1321
Db       1319 CAACTCCCGGAAGATTGAATGATCTGCAAATGAGTAACTTCGTCAATCTGAAGAATATAA 1378

Qy       1322 AATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGA 1381
Db       1379 CCTACTTGGTTTTAGATGAAGCAGACAAGATGTTGGACATGGGATTTGAACCCCAGATAA 1438

Qy       1382 AGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCA 1441
Db       1439 TGAAGATTTTGT------------TAGATGTGCGCCCAGATAGGCAGACAGTTATGACCA 1486

Qy       1442 GTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTT---AAAGTCAAATT 1498
Db       1487 GTGCTACATGGCCTCATTCAGTTCATCGCCTCGCACAATCTTATTTGAAAGAACCAATGA 1546

Qy       1499 ATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTC 1558
```

```
Db       1547 TTGTCTATGTTGGTACATTGGATCTAGTTGCTGTAAGTTCAGTGAAGCAAAATATAATTG 1606

Qy       1559 TCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGG 1618

Db       1607 TAACCACCGAGGAAGAGAAATGGAGTCACATGCAAACTTTTCTACAGAGTATGTCATCCA 1666

Qy       1619 ATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTC 1678

Db       1667 CAGACAAAGTCATTGTCTTCGTTTCTCGAAAAGCTGTTGCGGATCACTTATCAAGTGACC 1726

Qy       1679 TTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGG 1738

Db       1727 TAATACTTGGAAATATATCAGTAGAGTCTCTGCATGGAGATAGAGAACAGAGAGATCGGG 1786

Qy       1739 AGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAG 1798

Db       1787 AGAAAGCATTAGAGAACTTTAAAACAGGCAAAGTGAGAATACTAATTGCAACTGATCTAG 1846

Qy       1799 CTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTA 1858

Db       1847 CCTCTAGAGGACTTGATGTCCATGACGTTACACATGTCTATAATTTTGACTTTCCACGGA 1906

Qy       1859 CCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAG 1918

Db       1907 ATATTGAAGAATACGTACACCGAATAGGGCGCACGGGAAGAGCAGGGAGGACTGGTGTTT 1966

Qy       1919 CAATTTCCTTTTT 1931

Db       1967 CCATTACAACTTT 1979
```

RESULT 7
PCT-US96-05320A-894
; Sequence 894, Application PC/TUS9605320A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences
; APPLICANT: 9410 Key West Avenue
; APPLICANT: Rockville, MD 20850
; APPLICANT: United States of America
; APPLICANT: Johns Hopkins University
; APPLICANT: 720 Rutland Avenue
; APPLICANT: Baltimore, MD 21205
; APPLICANT: United States of America
; APPLICANT: Mark D. Adams
; APPLICANT: Owen White
; APPLICANT: Hamilton O. Smith
; APPLICANT: J. Craig Venter
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome,
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20003-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05320A
; FILING DATE: April22, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Eric K. Steffe
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.014PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 894:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
PCT-US96-05320A-894

Query Match 5.7%; Score 123.6; DB 5; Length 1254;
Best Local Similarity 47.1%; Pred. No. 6.1e-26;
Matches 495; Conservative 0; Mismatches 539; Indels 18; Gaps 3;

```
Qy        908 TTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATAC 967

Db         80 TGGCAAAAAAAGGCTTTGATTTTTGTACCCCAATTCAGGCTTTATCCTTGCCTATCAGTT 139

Qy        968 TTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTC 1027

Db        140 TAAATGGACGAGATGTCGCAGGACAAGCTCAAACTGGTACAGGCAAGACAATGGCTTTTT 199

Qy       1028 TCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGT 1087

Db        200 TAACGGCTACTTTTCACCATCTTTTAACTCACCAAGA---------TCCTAATCTTAAAT 250

Qy       1088 TGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATT 1147

Db        251 ATCCTCACCCAAGAGCTTTGATTTTAGCACCTACTCGAGAATTAGCGGTACAGATTAGTA 310

Qy       1148 TGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAA 1207

Db        311 ATGACGCAGAATTTCTTGCAAAAGCGAGTGGATTAAAGACCGCACTTGCCTATGGTGGCG 370

Qy       1208 CCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTC 1267

Db        371 ATGGTTATGATAAACAACTACAAGCGATTGAGCGTGGCGTCGATATTTTGATTGGTACGA 430

Qy       1268 CTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACT 1327

Db        431 CGGGGCGAGTCATTGATTATGTGAAACAAGGCGTAATTGGTTTAGATGAAATCCAAGTTG 490

Qy       1328 TAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGT 1387

Db        491 TCGTGCTAGATGAAGCAGATCGAATGTTTGATCTTGGGTTTATCCGTGATATTCGTTATT 550

Qy       1388 TAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAA 1447

Db        551 TATTGCGTAA------ATGCCCCGCTCCGCAAGCTCGTTTAACGATGTTATTTTCAGCGA 604

Qy       1448 CTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTG 1507

Db        605 CGCTTTCTTATAAAGTGCGTGAATTAGCATTTGAAGATATGAATGAACCTGA---ATATA 661

Qy       1508 TTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTG 1567

Db        662 TTGAAATTGAACCAGAACAAAAAACAGGGCACCGAATTAAAGAAGAACTTTTTATCCAT 721

Qy       1568 GCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAA 1627

Db        722 CTAATCAGGATAAAATGGCACTTCTCTTAACCTTAATGGAAGATGAATGGCCTGAACGCT 781

Qy       1628 CTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAG 1687

Db        782 GTATTGTATTTGCGAATACGAAACATCGTTGTGAAGAAATTTGGGGCTATTTGGCGGCTG 841

Qy       1688 AAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTC 1747

Db        842 ATGGGCATCGTGTCGGTTTACTGACTGGCGATGTAGCACAGAAAAAACGTTTATCGTTAT 901
```

```
Qy       1748 TTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAG 1807
Db        902 TAAAACAATTTACTGATGGTGATTTGGATATTTAGTGGCAACAGATGTGGCTGCTCGTG 961

Qy       1808 GGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATG 1867
Db        962 GCTTGCATATTTCTGATGTGACGCATGTTTTCAATTATGATTTACCCGATGATCGAGAAG 1021

Qy       1868 AATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCT 1927
Db       1022 ATTATGTTCACCGAATTGGGCGTACTGGACGAGCAGGGGAAAGTGGTGTTTCGATTAGTT 1081

Qy       1928 TTTTTGATCTTGAATCGGATAACCATTTAGCA 1959
Db       1082 TCGCTTGTGAAGAATATGCGATGAATTTACCA 1113
```

RESULT 8

```
US-08-961-527-203/c
; Sequence 203, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  203:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3230 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-527-203

  Query Match             5.4%;  Score 116.8;  DB 4;  Length 3230;
  Best Local Similarity   47.0%;  Pred. No. 9.5e-24;
  Matches  505;  Conservative    0;  Mismatches  527;  Indels   42;  Gaps    3;

Qy        863 TGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTT 922
Db       1910 TGAAATTTAATGAATTAAACTTGTCTGCTGATTTGCTAGCAGAAATTGAAAAAGCTGGTT 1851

Qy        923 ATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATT 982
Db       1850 TTGTAGAAGCTAGTCCTATCCAAGAACAAACTATTCCTTTGGCCCTTGAAGGCAAGGATG 1791

Qy        983 TGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGG 1042
Db       1790 TTATCGGTCAAGCTCAGACTGGTACAGGAAAAACTGCAGCCTTTGGCTTGCCTACCCTTG 1731

Qy       1043 CTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGT 1102
Db       1730 AAAAAATCCGTACAGAAGAAGCGACTATCCAAGCCTTGG-------------------- 1692

Qy       1103 GTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAAT 1162
Db       1691 ------TCATCGCTCCAACTCGTGAACTAGCTGTCCAAAGTCAAGAAGAACTCTTCCGCT 1638

Qy       1163 TTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATT 1222
Db       1637 TTGGTCGTAGTAAGGGAGTCAAAGTCCGTTCAGTATATGGCGGATCAAGCATTGAAAAAC 1578

Qy       1223 CAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGG 1282
Db       1577 AAATTAAGGCTCTTAAATCTGGTGCCCATATTGTGGTGGGAACTCCAGGTCGCCTCTTGG 1518

Qy       1283 ATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAG 1342
Db       1517 ACTTGATTAAACGCAAGGCCTTGAAATTACAAGACATTGAAACCCTTATCCTTGACGAAG 1458

Qy       1343 CTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAG 1402
Db       1457 CGGATGAAATGCTTAACATGGGCTTCCTTGAAGACATCGAAGCCATTATTTC-------- 1406

Qy       1403 GAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAA 1462
Db       1405 ----CCGTGTACCTGAGAACCGTCAAACTTTGCTTTTCTCAGCAACTATGCCAGATGCCA 1350

Qy       1463 TTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAG 1522
Db       1349 TCAAACGTATCGGTGTTCAGTTTATGAAAGCCCCTGAAC---ATGTCAAGATTGCGGCTA 1293

Qy       1523 TGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAA 1582
Db       1292 AGGAATTGACAACAGAATTGGTTGACCAGTACTATATCCGTGTTAAGGAACAAGAAAAAT 1233

Qy       1583 GAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTG 1642
Db       1232 TTGACACCATGACTCGTCTCATGGATGTGGCACAACCAGAACTCGCTATTGTATTTGGTC 1173

Qy       1643 AAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTA 1702
Db       1172 GTACCAAACGCCGTGTGGATGAATTGACTCGTGGTTTGAAAATTCGTGGCTTCCGTGCAG 1113

Qy       1703 CAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCT 1762
Db       1112 AAGGAATTCATGGCGACCTAGACCAAAACAAACGTCTTCGTGTCCTTCGTGACTTTAAAA 1053

Qy       1763 TTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAA 1822
Db       1052 ATGGCAATCTTGATGTTTTGGTTGCGACAGACGTTGCAGCGCGTGGTTTGGATATTTCAG 993

Qy       1823 ATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAA 1882
Db        992 GTGTGACCCATGTCTACAACTACGATATTCCACAAGATCCTGAGAGTTATGTTCACCGTA 933

Qy       1883 TTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATC 1936
Db        932 TCGGTCGTACAGGTCGTGCTGGTAAGTCAGGTCAATCTATTACTTTGTTGCTC 879
```

RESULT 9

```
US-09-208-742-3
; Sequence 3, Application US/09208742
; Patent No. 6174679
; GENERAL INFORMATION:
;  APPLICANT: Kaufmann, Joerg
;  TITLE OF INVENTION: CIF150/hTAFII150 is Necessary for Cell
;  TITLE OF INVENTION: Cycle Progression
;  FILE REFERENCE: 1453.002
```

```
;  CURRENT APPLICATION NUMBER: US/09/208,742
;  CURRENT FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 3825
;   TYPE: DNA
;   ORGANISM: human
US-09-208-742-3

  Query Match             5.3%;  Score 115.2;  DB 4;  Length 3825;
  Best Local Similarity   46.3%;  Pred. No. 3e-23;
  Matches  513;  Conservative    0;  Mismatches  568;  Indels   27;  Gaps    3;

Qy        832 GTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAG 891
Db        862 GTCTCTGGTGCTGCACCTCCTAGACCAGGAAGTAGCTTTGCTCATTTTGGGTTTGACGAA 921

Qy        892 ACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATAC 951
Db        922 CAACTTATGCACCAGATTCGGAAATCTGAATACACACAGCCCACTCCAATACAGTGCCAG 981

Qy        952 AGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGG 1011
Db        982 GGTGTGCCTGTGGCATTAAGTGGTAGAGACATGATTGGTATTGCCAAAACAGGTAGTGGG 1041

Qy       1012 AAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCC 1071
Db       1042 AAAACTGCAGCCTTCATTTGGCCCATGTTGATTCATATAATGGACCA-----------G 1089

Qy       1072 AGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTG 1131
Db       1090 AAGGAGTTGGAACCAGGTGATGGACCAATTGCAGTGATTGTGTGTCCTACCAGGGAGCTT 1149

Qy       1132 GTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTT 1191
Db       1150 TGCCAGCAGATCCATCCAGAATGTAAGCGGTTTGGAAAAGCATATAATCTTCGATCAGTG 1209

Qy       1192 GTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAAT 1251
Db       1210 GCCGTATATGGAGGAGGGAGTATGTGGGAGCAGGCCAAGGCCCTTCAGGAGGGGGCAGAG 1269

Qy       1252 ATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTC 1311
Db       1270 ATTGTTGTGTGTACCCCAGGTCGACTGATAGATCATGTGAAAAAGAAAGCTACCAATCTT 1329

Qy       1312 AAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGT 1371
Db       1330 CAAAGAGTCTCTTACCTTGTGTTTGATGAAGCAGATCGAATGTTTGACATGGGATTTG-- 1387

Qy       1372 CCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACC 1431
Db       1388 ----------AGTACCAAGTTCGATCCATAGCAAGTCATGTTCGTCCTGACAGGCAGACT 1437

Qy       1432 CTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAG 1491
Db       1438 CTCTTATTTAGTGCAACTTTTCGGAAGAAGATTGAAAAGTTGGCCAGAGACATCCTGATC 1497

Qy       1492 TCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAG 1551
Db       1498 GACCCTATTCGAGTGGTGCAGGGAGATATTGGAGAGGCAAATGAAGATGTGACACAGATT 1557

Qy       1552 ACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAAC 1611
Db       1558 GTGGAGATTCTCCATTCTGGACCTAGTAAATGGAACTGGCTTACCCGGCGTCTGGTAGAA 1617

Qy       1612 ATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAA---AAAGCAGATTTTACT 1668
Db       1618 TTTACCTCTTCAGGGAGTGTCCTCCTCTTTGTTACTAAAAAAGCCAATGCTGAAGAGCTA 1677

Qy       1669 GCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAG 1728
Db       1678 GCGAATAACCTTAAACAGGAGGGTCATAATCTTGGGCTGCTCCATGGGGATATGGATCAG 1737

Qy       1729 AGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCT 1788
Db       1738 AGTGAGAGAAACAAGGTCATTTCAGACTTTAAGAAAAAGGACATCCCAGTCCTGGTGGCC 1797

Qy       1789 ACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGAT 1848
Db       1798 ACAGATGTTGCAGCCCGTGGTCTGGACATTCCTTCAATTAAGACTGTCATTAACTATGAT 1857

Qy       1849 CTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAAT 1908
Db       1858 GTGGCACGGGACATTGATACGCACACTCACAGGATTGGCCGCACAGGAAGAGCGGGTGAG 1917

Qy       1909 ACTGGCAGAGCAATTTCCTTTTTGATC 1936
Db       1918 AAAGGTGTGGCCTATACCCTACTCACTC 1945
```

```
RESULT 10
US-09-134-001C-1425
; Sequence 1425, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1425
;   LENGTH: 1587
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1425

  Query Match             4.8%;  Score 103.2;  DB 4;  Length 1587;
  Best Local Similarity   48.5%;  Pred. No. 5.7e-20;
  Matches  411;  Conservative    0;  Mismatches  423;  Indels   14;  Gaps    4;

Qy       1090 CAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTG 1149
Db        259 CAGGGTGTTCAATCATTAATTCTTGCGCCAACGAGAGAGCTTGCTATGCAAGTAGCTGAG 318

Qy       1150 GAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACC 1209
Db        319 CAATTAAGAGAGTTCAGTAAAGGTCAAAAAGTACAAGTTGTTACAGTCTTTGGTGGTATG 378

Qy       1210 CAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCT 1269
Db        379 CCTATTGAAAGACAGATTAAAGCATTAAAAAGAGGCCCTCAAATTGTAGTAGGTACGCCG 438

Qy       1270 GGAAGACTGATGGAT-ATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTT 1328
Db        439 GGAAGAGTTATTGATCATCTTAATCGTCGTACACTTAAAACTCAAGGAATTCACACGCTT 498

Qy       1329 AGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTT 1388
Db        499 A-TTTTAGATGAAGCAGATGAAATGATGAATATGGGATTCATCGATGATATGAGATTTAT 557

Qy       1389 AATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAAC 1448
Db        558 TAT---------GGATAAAATTCCAGCTGAACAACGTCAAACTATGCTATTTTCTGCTAC 608

Qy       1449 TTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGT 1508
Db        609 GATGCCAAAAGCAATCCAAGAATTGGTACAGCAGTTTATGAAAGCTCCTAAAATAATTAA 668

Qy       1509 TGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGG 1568
```

```
Db        669 GACAATGAATAATGAAATGTCAGATCCTCAAATTGATGAATATTATACAATC---GTTAA 725

Qy       1569 CCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAAC 1628
Db        726 AGAATTAGAGAAATTTGATACATTTACAAACTTTTTAGATGTTCATCAGCCTGAATTGGC 785

Qy       1629 TATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGA 1688
Db        786 GATTGTCTTTGGTCGTACTAAGCGTCGCGTAGATGAATTAACAAGCGCGCTATTATCTAA 845

Qy       1689 AAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCT 1748
Db        846 AGGATACAAAGCAGAAGGTTTACATGGGGATATTACCCAAGCAAAACGTTTAGAAGTATT 905

Qy       1749 TGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGG 1808
Db        906 GAAGAAATTTAAGAACGATCAAATAGATATTTTAGTTGCTACAGATGTGGCTGCACGTGG 965

Qy       1809 GCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGA 1868
Db        966 TCTTGATATTTCTGGTGTGAGTCATGTTTATAATTTTGATATCCCTCAAGATACAGAAAG 1025

Qy       1869 ATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTT 1928
Db       1026 TTATACGCACCGTATTGGACGTACTGGTCGAGCAGGAAAAGAAGGTATTGCAGTTACTTT 1085

Qy       1929 TTTTGATC 1936
Db       1086 TGTAAATC 1093
```

RESULT 11
US-09-318-443-1
; Sequence 1, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318,443
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1245
;   TYPE: DNA
;   ORGANISM: Xenopus laevis
US-09-318-443-1

```
  Query Match             4.2%;  Score 92.2;  DB 4;  Length 1245;
  Best Local Similarity   46.0%;  Pred. No. 7.7e-17;
  Matches  396;  Conservative    0;  Mismatches  453;  Indels   12;  Gaps    2;

Qy       1080 TAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCA 1139
Db        312 TATTCAGATCCGTGAAACCCAAGCCTTGATTTTAGCACCCACCAAAGAGTTAGCACGGCA 371

Qy       1140 GATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATA 1199
Db        372 AATTCAGAAGGTGTTGCTTGCTTTGGGGGACTACATGAATGTGCAGTGTCATGCGTGTAT 431

Qy       1200 TGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATG 1259
Db        432 TGGAGGCACAAATGTTGGAGAGGATATCCGAAAATTGGATTATGGGCAGCACGTTGTTGC 491

Qy       1260 TGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGAT 1319
Db        492 TGGAACACCAGGGCGTGTTTTTGATATGATTCGACGCAGAAGTTTAAGAACTCGGGCCAT 551

Qy       1320 CAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAAT 1379
Db        552 TAAAATGTTAGTGCTGGATGAAGCTGATGAAATGTTGAATAAGGGTTTC--------AAG 603

Qy       1380 GAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTT 1439
Db        604 GAGCAAATTTATGATGTATACAGGTATCTGCCTCCAGCAACA----CAAGTTTGTTTAAT 659

Qy       1440 CAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTA 1499
Db        660 CAGTGCTACCCTGCCACATGAAATCCTGGAAATGACCAATAAGTTTATGACTGATCCCAT 719

Qy       1500 TCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCT 1559
Db        720 CCGTATCCTTGTGAAACGTGATGAGTTGACACTGGAAGGCATCAAGCAGTTTTTTGTGGC 779

Qy       1560 CCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGA 1619
Db        780 AGTGGAGAGAGAAGAGTGGAAATTTGATACTTTGTGTGATTTATATGACACTTTGACTAT 839

Qy       1620 TGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCT 1679
Db        840 TACACAAGCTGTAATCTTCTGCAACACCAAAAGAAAGGTAGATTGGTTGACTGAAAAAAT 899

Qy       1680 TTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGA 1739
Db        900 GAGAGAAGCAAATTTCACAGTTTCGTCAATGCATGGTGATATGCCCCAAAAGGAGAGAGA 959

Qy       1740 GCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGC 1799
Db        960 GTCAATCATGAAAGAATTCCGATCTGGTGCAAGCCGAGTCCTCATATCAACGGACGTCTG 1019

Qy       1800 TGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTAC 1859
Db       1020 GGCCCGAGGATTGGATGTGCCACAGGTCTCCTTGATTATCAACTATGATCTTCCCAATAA 1079

Qy       1860 CATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGC 1919
Db       1080 CCGAGAATTGTACATTCACAGAATTGGCCGATCAGGAAGATATGGAAGAAAGGGTGTTGC 1139

Qy       1920 AATTTCCTTTTTTGATCTTGA 1940
Db       1140 CATTAACTTTGTCAAGAATGA 1160
```

RESULT 12
US-09-134-001C-816
; Sequence 816, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 816
;   LENGTH: 1368
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-816

```
  Query Match             4.2%;  Score 91.2;  DB 4;  Length 1368;
  Best Local Similarity   45.6%;  Pred. No. 1.6e-16;
  Matches  523;  Conservative    0;  Mismatches  583;  Indels   42;  Gaps    4;

Qy        857 CAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAG 916
Db         26 CAAAACATCCATTTGAACACTTTAATTTAGATGAGAATTTAATTGAAGCTGTTAAAAATC 85
```

```
Qy         917 CTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGAC 976
                 || |     || |  |||     ||||||    |  || || |   || |||  ||||
Db          86 TCAATTTTGAAAAACCGACTGAAATCCAAAATAGAATCATACCGAGAATTCTTAAAGGAA 145

Qy         977 GAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAA 1036
                | |||| || |       |||||| ||  |||| ||| |    |||||||| |||||
Db         146 CAAATTTAATAGGACAATCTCAAACTGGAACTGGAAAGTCACACGCTTTTCTTTTACCAT 205

Qy        1037 TTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAAC 1096
               |  |    | |||       |||                            | || || |
Db         206 TAATTCAACTTATAGAAAGTGACA--------------------------TTCAAGAGC 238

Qy        1097 CAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTGGAAGCCA 1156
               || |     ||   | ||||| ||||| || ||| | |    | ||   | |||    | || |
Db         239 CACAAGCCATCGTAGTAGCTCCAACACGTGAACTTGCTCAGCAACTATATCAAGTTGCTA 298

Qy        1157 GAAAATTTTCT---TTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGC 1213
                 | ||        |     | |  | | |||     |     ||  | ||| || |||  |
Db         299 TGCATTTAGTAAAATTCAAAAAAGGTATAAATGTAAAACTTTTCATTGGTGGTACCGATT 358

Qy        1214 TGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAA 1273
               | | | |      | | || |       ||          |   |   |  || ||| ||   ||
Db         359 TAGAAAAAGATAAACAACGATGTAGCCATCAACCACAACTCATTATTGGTACACCAACAA 418

Qy        1274 GACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTT 1333
               || |  |  |||  |     |      ||    |  ||   |      |      ||| |||  ||
Db         419 GAATTAATGATTTAGCACATTCAGGTTATCTTCATGCACATTTAGCGTCATATTTAATTA 478

Qy        1334 TGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTT 1393
               |  ||||||||||||   |||  | ||   | |||  |    |  |||   |   |   |
Db         479 TAGATGAAGCTGATTTAATGATTGACCTCGGTCTCATTGAAGATGTTGACCA-------- 530

Qy        1394 CTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTC 1453
                  |   |||  |    |      | ||| |     ||          | || ||||||||   | |
Db         531 -TATTGCAGCGAGATTAGATGATGAAAATGCTCATCTAGCGGTATTTAGTGCAACAATAC 589

Qy        1454 CAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTG 1513
               |  |   |  | | |||  ||        | |  |||||     |     |  |  |  |  | ||
Db         590 CTAAATCATTACAACCATTTTTAAATAAATATTTAAGTCAACCAGAATTTGTAGAAGTTG 649

Qy        1514 TTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGT 1573
                ||| |||   |||     |      |     |      |    |          |||    |   ||   |
Db         650 ATGGCAAAGCTCATAATAAAGAAAATATCGAATTTTATCTAATTCCTACAAAAGG---TT 706

Qy        1574 TCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGGATGAAAGAACTATGG 1633
                    | ||    ||| ||       | ||| |   |  |    |        |  |    |  |||
Db         707 CTGCTAAGGTAGATAAAACATTGGAATTGATAGATATATTGAATCCTTATCTATGTATTA 766

Qy        1634 TCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAA 1693
               | ||     | |  | |      |||| || ||      |||     | |  |||||  | |
Db         767 TTTTCTGTAACAGTCGTGAAAATGCCGATGAATTGGCAGACACTTTAAATAAAGAAGGAA 826

Qy        1694 TATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAG 1753
               |     ||       |    | ||||||||        ||   ||||| ||   | |||    |    |
Db         827 TTAAAATAGGTATGATTCATGGTGGTTTAACACCAAGAGAACGTAAACAACAAATGAAAA 886

Qy        1754 ATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGG 1813
                   |   |    |  | | |  | ||  | ||  ||||  |        |  |||  |  |||||  | |
Db         887 GAATAAGAAATTTAGATTTTCAATTTGTCATTGCAAGCGATCTTGCTTCTAGAGGAATAG 946

Qy        1814 ATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATG 1873
               ||||||||     ||      |||||||| |||||  ||| |  ||      |    || |||     |
Db         947 ATATTGAAGGCGTAAGTCATGTTATTAATTTCGATGTACCCAATGATATCGATTTCTTCA 1006

Qy        1874 TTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTG 1933
                   |||||  | || || ||  ||  ||  ||      |||  ||   ||| |||  | |||   |
Db        1007 CACATCGCGTAGGTCGAACAGGAAGAGGTAATTATAAAGGTGTAGCCATTACATTATATA 1066

Qy        1934 ATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAAC 1993
                 || |||      || |    || |  |  |      |   |  ||  |  |  |  ||   |
Db        1067 GTCCTGATGAAGAAAGTAATATTACTCTTATTGAAGACAGAGGTTATAAATTTGAAAATG 1126

Qy        1994 AGGATGTT 2001
                 ||| ||
Db        1127 TAGATATT 1134
```

RESULT 13
US-09-222-575-126
; Sequence 126, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/09/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 126
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Human
US-09-222-575-126

```
  Query Match             4.2%;  Score 90.8;  DB 4;  Length 490;
  Best Local Similarity   53.4%;  Pred. No. 1.2e-16;
  Matches  222;  Conservative    0;  Mismatches  182;  Indels   12;  Gaps    1;

Qy         973 GGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTA 1032
                |||    ||| || |  |         || || || ||  |||||||| ||   | |||  | |  ||
Db          34 GGATTGGATATGGTTGGAGTGGCACAGACTGGATCTGGGAAAACATTGTCTTATTTGCTT 93

Qy        1033 CCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAG 1092
                ||     | |  || || |  |||  |     | |      |  ||  |                |
Db          94 CCTGCCATTGTCCACATCAATCATCAGCCATTCCTAGAGAGAGG-----------CGAT 141

Qy        1093 GAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTGGAA 1152
                  | ||    |    |   |||||||||||| ||| |  | | ||| |   |        | |
Db         142 GGGCCTATTTGTTTGGTGCTGGCACCAACTCGGGAACTGGCCCAACAGGTGCAGCAAGTA 201

Qy        1153 GCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAG 1212
                ||     ||| || |    |      |  |  |   |    ||    ||| || || ||  | |
Db         202 GCTGCTGAATATTGTAGAGCATGTCGCTTGAAGTCTACTTGTATCTACGGTGGTGCTCCT 261

Qy        1213 CTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGA 1272
                    |||||     ||| ||    |   |       |||        | || |       ||| || ||||||
Db         262 AAGGGACCACAAATACGTGATTTGGAGAGAGGTGTGGAAATCTGTATTGCAACACCTGGA 321

Qy        1273 AGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTT 1332
               |||||||| ||    |   |||         | |||  |        ||| |  |      |   |   ||| |  ||
Db         322 AGACTGATTGACTTTTTAGAGTGTGGAAAAACCAATCTGAGAAGAACAACCTACCTTGTC 381

Qy        1333 TTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTT 1388
                 |||||||||| |||  |  |||  |  |||||||| ||||   ||   ||||  |  ||||  |
Db         382 CTTGATGAAGCAGATAGAATGCTTGATATGGGCTTTGAACCCCAAATAAGGAAGAT 437
```

RESULT 14
US-08-959-749-1
; Sequence 1, Application US/08959749
; Patent No. 5989858
; GENERAL INFORMATION:
; APPLICANT: Huang, Jianzhong
; APPLICANT: McDevitt, Damien
; APPLICANT: Wang, Min
; APPLICANT: Traini, Christopher M.
; TITLE OF INVENTION: No. 5989858el DbpB
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads

```
;      STREET:  4000 Bell Atlantic Tower, 1717 Arch Stre
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2793
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/959,749
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dickinson, Q. Todd
;      REGISTRATION NUMBER:  28,354
;      REFERENCE/DOCKET NUMBER:  GM10113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215/994-2252
;      TELEFAX:  215/994-2222
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1347 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-959-749-1

  Query Match              3.9%;  Score 84.6;  DB 2;  Length 1347;
  Best Local Similarity   53.5%;  Pred. No. 1.3e-14;
  Matches  177;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy       1629 TATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGA 1688

Db        738 TATTATTTTCTGTAATAGTAGAGATAATGCAAATGATTTAGCACGTTCACTAAATGAAGC 797

Qy       1689 AAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCT 1748

Db        798 TGGTATTAAAGTTGGTATGATTCATGGTGGTTTAACGCCACGTGAACGTAAACAACAAAT 857

Qy       1749 TGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGG 1808

Db        858 GAAACGTATACGTAATTTAGAATTCCAATACGTTATTGCCAGCGATTTAGCATCTCGTGG 917

Qy       1809 GCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGA 1868

Db        918 TATTGATATTGAAGGTGTTAGTCATGTCATCAATTTTGATGTGCCAAATGATATTGACTT 977

Qy       1869 ATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTT 1928

Db        978 CTTTACGCATAGAGTTGGACGAACTGGTCGTGGGAATTATAAAGGTGTAGCAATTACGCT 1037

Qy       1929 TTTTGATCTTGAATCGGATAACCATTTAGCA 1959

Db       1038 TTATAGTCCTGATGAAGAACACAATATTTCA 1068

RESULT 15
US-09-351-497-1
; Sequence 1, Application US/09351497
; Patent No. 6313270
;  GENERAL INFORMATION:
;    APPLICANT:  Huang, Jianzhong
;    APPLICANT:  McDevitt, Damien
;    APPLICANT:  Wang, Min
;    APPLICANT:  Traini, Christopher M.
;    TITLE OF INVENTION:  No. 6313270el DbpB
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert, Price & Rhoads
;      STREET:  4000 Bell Atlantic Tower, 1717 Arch Stre
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2793
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/351,497
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/959,749
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dickinson, Q. Todd
;      REGISTRATION NUMBER:  28,354
;      REFERENCE/DOCKET NUMBER:  GM10113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215/994-2252
;      TELEFAX:  215/994-2222
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1347 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-351-497-1

  Query Match              3.9%;  Score 84.6;  DB 4;  Length 1347;
  Best Local Similarity   53.5%;  Pred. No. 1.3e-14;
  Matches  177;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy       1629 TATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGA 1688

Db        738 TATTATTTTCTGTAATAGTAGAGATAATGCAAATGATTTAGCACGTTCACTAAATGAAGC 797

Qy       1689 AAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCT 1748

Db        798 TGGTATTAAAGTTGGTATGATTCATGGTGGTTTAACGCCACGTGAACGTAAACAACAAAT 857

Qy       1749 TGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGG 1808

Db        858 GAAACGTATACGTAATTTAGAATTCCAATACGTTATTGCCAGCGATTTAGCATCTCGTGG 917

Qy       1809 GCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGA 1868

Db        918 TATTGATATTGAAGGTGTTAGTCATGTCATCAATTTTGATGTGCCAAATGATATTGACTT 977

Qy       1869 ATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTT 1928

Db        978 CTTTACGCATAGAGTTGGACGAACTGGTCGTGGGAATTATAAAGGTGTAGCAATTACGCT 1037

Qy       1929 TTTTGATCTTGAATCGGATAACCATTTAGCA 1959

Db       1038 TTATAGTCCTGATGAAGAACACAATATTTCA 1068

Search completed: June 10, 2003, 10:42:46
Job time : 142 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:         June 10, 2003, 09:43:39 ; Search time 296 Seconds
                                          (without alignments)
                                          10269.578 Million cell updates/sec

Title:          US-09-714-865-15
Perfect score:  2172
Sequence:       1 atgggggatgaagattggga..........tagatgatgagtcatgggat 2172

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters:      1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                 1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                 2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                 3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                 4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                 5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                 6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                 7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                 8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                 9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                 10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                 11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                 12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                 13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                 14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	389.8	17.9	420	10	US-09-833-381-1475	Sequence 1475, Ap
	2	386.4	17.8	2451	12	US-10-044-090-480	Sequence 480, App
	3	384.8	17.7	4423	12	US-10-002-600-67	Sequence 67, Appl
	4	317	14.6	591	9	US-09-764-891-1170	Sequence 1170, Ap
	5	275.8	12.7	1902	9	US-09-938-842A-521	Sequence 521, App
	6	220	10.1	2194	9	US-09-764-891-7481	Sequence 7481, Ap
	7	220	10.1	2195	9	US-09-764-891-7482	Sequence 7482, Ap
	8	220	10.1	2197	9	US-09-764-891-7483	Sequence 7483, Ap
	9	210.2	9.7	3841	9	US-10-198-846-12716	Sequence 12716, A
	10	168.4	7.8	403	10	US-09-960-352-14381	Sequence 14381, A
	11	166.6	7.7	2010	10	US-09-964-824A-268	Sequence 268, App
	12	166.6	7.7	2010	10	US-09-880-107-3397	Sequence 3397, Ap
	13	164.4	7.6	2247	9	US-09-938-842A-1091	Sequence 1091, Ap
	14	159.6	7.3	3501	9	US-09-938-842A-13	Sequence 13, Appl
	15	158.2	7.3	2365	10	US-09-923-831-42	Sequence 42, Appl
	16	153.6	7.1	1614	9	US-09-938-842A-2560	Sequence 2560, Ap
	17	145.2	6.7	418	9	US-09-918-995-35717	Sequence 35717, A
	18	136.6	6.3	1052	10	US-09-925-301-597	Sequence 597, App
	19	136	6.3	1860	9	US-09-938-842A-699	Sequence 699, App
c	20	125.6	5.8	1830121	9	US-10-329-960-1	Sequence 1, Appli
	21	123.6	5.7	1830121	9	US-10-329-960-1	Sequence 1, Appli
	22	120	5.5	2202	9	US-09-938-842A-2296	Sequence 2296, Ap
	23	118.4	5.5	1575	10	US-09-815-242-9393	Sequence 9393, Ap
	24	117	5.4	1833	9	US-09-938-842A-1839	Sequence 1839, Ap
	25	115.2	5.3	3347	9	US-09-736-457-318	Sequence 318, App
	26	115.2	5.3	3347	9	US-09-902-941-318	Sequence 318, App
	27	115.2	5.3	3347	9	US-09-849-626-318	Sequence 318, App
	28	115.2	5.3	3347	9	US-10-017-754-318	Sequence 318, App
	29	112.6	5.2	2936	9	US-10-108-605-54	Sequence 54, Appl
	30	109.6	5.0	516	10	US-09-920-300A-784	Sequence 784, App
	31	109.6	5.0	516	12	US-10-033-528-784	Sequence 784, App
c	32	109.6	5.0	551	10	US-09-920-300A-1344	Sequence 1344, Ap
c	33	109.6	5.0	551	12	US-10-033-528-1344	Sequence 1344, Ap
c	34	109.6	5.0	573	9	US-09-736-457-1384	Sequence 1384, Ap
c	35	109.6	5.0	573	9	US-09-902-941-1384	Sequence 1384, Ap
	36	109.6	5.0	573	9	US-09-796-692-245	Sequence 245, App
	37	109.6	5.0	573	9	US-09-796-692-375	Sequence 375, App
	38	109.6	5.0	573	9	US-09-796-692-397	Sequence 397, App
	39	109.6	5.0	573	9	US-09-796-692-492	Sequence 492, App
	40	109.6	5.0	573	9	US-09-796-692-503	Sequence 503, App
	41	109.6	5.0	573	9	US-09-796-692-5064	Sequence 5064, Ap
	42	109.6	5.0	573	9	US-09-796-692-5261	Sequence 5261, Ap
c	43	109.6	5.0	573	9	US-09-849-626-1384	Sequence 1384, Ap
c	44	109.6	5.0	573	9	US-10-017-754-1384	Sequence 1384, Ap
	45	109.6	5.0	573	9	US-10-040-862-245	Sequence 245, App

ALIGNMENTS

```
RESULT 1
US-09-833-381-1475
; Sequence 1475, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1475
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-1475

  Query Match             17.9%;  Score 389.8;  DB 10;  Length 420;
  Best Local Similarity   99.5%;  Pred. No. 2.2e-100;
  Matches  391;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        805 GACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTG 864
              | |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         28 GTCAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAGTTCTG 87

Qy        865 ACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTAT 924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 ACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTAT 147

Qy        925 ACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTG 984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTG 207

Qy        985 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCT 1044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 ATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCT 267

Qy       1045 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 1104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        268 CATATGATGCATGATGGAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGT 327

Qy       1105 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 1164

Db        328 ATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTT 387

Qy       1165 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATA 1197

Db        388 TCTTTTGGGACTTGTGTAAGAGCTGTTGTTATA 420
```

```
RESULT 2
US-10-044-090-480
; Sequence 480, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 480
;   LENGTH: 2451
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 035282CB1
US-10-044-090-480

  Query Match             17.8%;  Score 386.4;  DB 12;  Length 2451;
  Best Local Similarity   58.7%;  Pred. No. 6.6e-99;
  Matches  763;  Conservative    0;  Mismatches  501;  Indels   36;  Gaps    4;

Qy        788 AGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATG 847

Db        503 ACACGGGGATTAACTTTGAGAAATATGATGATATACCAGTAGAGGCAACCGGCAGTAACT 562

Qy        848 CACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACA 907

Db        563 GTCCTCCACATATTGAGAATTTTAGCGATATTGACATGGGAGAAATTATCATGGGGAACA 622

Qy        908 TTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATAC 967

Db        623 TTGAACTTACTCGCTATACTCGTCCTACTCCAGTGCAAAAACATGCCATTCCTATTATTA 682

Qy        968 TTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTC 1027

Db        683 AGGGAAAAAGAGACTTAATGGCTTGTGCCCAAACAGGATCTGGGAAAACTGCAGCATTTC 742

Qy       1028 TCCTACCAATTTTGGCTCA--------------------------TATGATGCATGATG 1060

Db        743 TTTTACCCATACTGAGTCAGATATATACAGATGGTCCAGGAGAAGCTTTGAAGGCTGTGA 802

Qy       1061 GAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAA 1120

Db        803 AGGTAAATGGAAGGTATGGGCGCCGCAAACAATATCCAATATCCTTGGTTTTAGCCCCAA 862

Qy       1121 CTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTG 1180

Db        863 CAAGAGAATTGGCTGTACAGATCTATGAGGAAGCCAGAAAATTTTCCTACCGATCTAGAG 922

Qy       1181 TAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTAC 1240

Db        923 TTCGTCCTTGTGTAGTTTATGGTGGTGCTGATATTGGTCAGCAGATTCGGGACTTAGAAC 982

Qy       1241 AAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAA 1300

Db        983 GTGGATGCCACTTGTTAGTAGCCACTCCAGGACGTCTAGTGGATATGATGGAAAGAGGAA 1042

Qy       1301 AGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATGTTGGATA 1360

Db       1043 AGATTGGATTAGACTTCTGCAAGTACTTAGTGTTGGATGAAGCTGATAGGATGCTGGATA 1102

Qy       1361 TGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAAC 1420

Db       1103 TGGGATTTGAACCTCAGATACGTCGTATAGTTGAACAAGATACTATGCCACCAAAGGGCG 1162

Qy       1421 AGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAG 1480

Db       1163 TTCGTCACACCATGATGTTTAGTGCTACTTTTCCTAAGGAAATACAGATGCTTGCTCGTG 1222

Qy       1481 AGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAG 1540

Db       1223 ACTTTTTGGATGAA---TATATCTTTTTGGCTGTAGGCAGAGTAGGCTCTACCTCTGAGA 1279

Qy       1541 ATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAA 1600

Db       1280 ACATCACACAGAAAGTAGTTTGGGTGGAAGACTTAGATAAACGGTCATTTCTACTGGACA 1339

Qy       1601 TTCTGCGAAACATAGGG---GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAG 1657

Db       1340 TTTTAGGTGCAACAGGGAGTGATTCACTTACTTTAGTGTTTGTGGAGACCAAAAAGGGAG 1399

Qy       1658 CAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTG 1717

Db       1400 CAGATTCCCTGGAGGATTTCTTATACCATGAAGGATATGCTTGTACTAGTATTCATGGAG 1459

Qy       1718 ATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTCGCTTTGGAAAGTGCCCAG 1777

Db       1460 ACCGGTCACAGAGAGATCGAGAGGAGGCCCTTCACCAGTTTCGCTCAGGAAAAAGCCCAA 1519

Qy       1778 TTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTA 1837

Db       1520 TTCTAGTGGCTACAGCTGTGGCAGCACGAGGACTAGACATTTCAAATGTGAGACATGTTA 1579

Qy       1838 TCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTC 1897

Db       1580 TCAATTTTGATTTGCCAAGTGATATTGAAGAATATGTGCATCGTATTGGCCGTACAGGAC 1639

Qy       1898 GTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAG 1957

Db       1640 GTGTAGGAAACCTGGGCCTTGCCACCTCATTCTTTAA---TGAAAAAAATATGAATATTA 1696

Qy       1958 CACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAG 2017

Db       1697 CAAAGGATTTGTTGGATCTTCTTGTAGAAGCTAAACAAGAAGTGCCTTCTTGGTTGGAAA 1756

Qy       2018 AAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAG 2057

Db       1757 ATATGGCTTATGAACACCACTACAAGGGTGGCAGTCGTGG 1796
```

```
RESULT 3
US-10-002-600-67
; Sequence 67, Application US/10002600
; Patent No. US20020137077A1
; GENERAL INFORMATION:
;  APPLICANT: Hopkins, Christopher M.
;  APPLICANT:  Peterson, David P.
;  APPLICANT:  Cocks, Benjamin G.
;  APPLICANT:  Hawkins, Phillip R.
;  TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS
;  FILE REFERENCE: PA-0042 US
;  CURRENT APPLICATION NUMBER: US/10/002,600
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: 60/243,521
;  PRIOR FILING DATE: 2000-10-25
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: PERL Program
; SEQ ID NO 67
;   LENGTH: 4423
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;  FEATURE:
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Template ID: 350667.1
US-10-002-600-67

  Query Match            17.7%;  Score 384.8;  DB 12;  Length 4423;
  Best Local Similarity  58.6%;  Pred. No. 2.8e-98;
  Matches  762;  Conservative    0;  Mismatches  502;  Indels   36;  Gaps    4;

Qy        788 AGACAGGCATAAACTTCGACAAATACGACACTATTCTTGTGGAAGTGTCTGGACATGATG 847
Db        536 ACACGGGGATTAACTTTGAGAAATATGATGATATACCAGTAGAGGCAACCGGCAGTAACT 595

Qy        848 CACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTGAATAACAACA 907
Db        596 GTCCTCCACATATTGAGAATTTTAGCGATATTGACATGGGAGAAATTATCATGGGGAACA 655

Qy        908 TTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTATCATAC 967
Db        656 TTGAACTTACTCGCTATACTCGTCCTACTCCAGTGCAAAAACATGCCATTCCTATTATTA 715

Qy        968 TTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGGCTTTTC 1027
Db        716 AGGGAAAAAGAGACTTAATGGCTTGTGCCCAAACAGGATCTGGGAAAACTGCAGCATTTC 775

Qy       1028 TCCTACCAATTTTGGCTCA--------------------------TATGATGCATGATG 1060
Db        776 TTTTACCCATACTGAGTCAGATATATACAGATGGTCCAGGAGAAGCTTTGAAGGCTGTGA 835

Qy       1061 GAATAACTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAA 1120
Db        836 AGGAAAATGGAAGGTATGGGCGCCGCAAACAATATCCAATATCCTTGGTTTTAGCCCCAA 895

Qy       1121 CTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTG 1180
Db        896 CAAGAGAATTGGCTGTACAGATCTATGAGGAAGCCAGAAAATTTTCCTACCGATCTAGAG 955

Qy       1181 TAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTAC 1240
Db        956 TTCGTCCTTGTGTAGTTTATGGTGGTGCTGATATTGGTCAGCAGATTCGGGACTTAGAAC 1015

Qy       1241 AAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAA 1300
Db       1016 GTGGATGCCACTTGTTAGTAGCCACTCCAGGACGTCTAGTGGATATGATGGAAAGAGGAA 1075

Qy       1301 AGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATA 1360
Db       1076 AGATTGGATTAGACTTCTGCAAGTACTTAGTGTTGGATGAAGCTGATAGGATGCTGGATA 1135

Qy       1361 TGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAAC 1420
Db       1136 TGGGATTTGAACCTCAGATACGTCGTATAGTTGAACAAGATACTATGCCACCAAAGGGCG 1195

Qy       1421 AGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAG 1480
Db       1196 TTCGTCACACCATGATGTTTAGTGCTACTTTTCCTAAGGAAATACAGATGCTTGCTCGTG 1255

Qy       1481 AGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAG 1540
Db       1256 ACTTTTTGGATGAA---TATATCTTTTTGGCTGTAGGCAGAGTAGGCTCTACCTCTGAGA 1312

Qy       1541 ATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAA 1600
Db       1313 ACATCACACAGAAAGTAGTTTGGGTGGAAGACTTAGATAAACGGTCATTTCTACTGGACA 1372

Qy       1601 TTCTGCGAAACATAGGG---GATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAG 1657
Db       1373 TTTTAGGTGCAACAGGGAGTGATTCACTTACTTTAGTGTTTGTGGAGACCAAAAAGGGAG 1432

Qy       1658 CAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTG 1717
Db       1433 CAGATTCCCTGGAGGATTTCTTATACCATGAAGGATATGCTTGTACTAGTATTCATGGAG 1492

Qy       1718 ATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAG 1777
Db       1493 ACCGGTCACAGAGAGATCGAGAGGAGGCCCTTCACCAGTTTCGCTCAGGAAAAAGCCCAA 1552

Qy       1778 TTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTA 1837
Db       1553 TTCTAGTGGCTACAGCTGTGGCAGCACGAGGACTAGACATTTCAAATGTGAGACATGTTA 1612

Qy       1838 TCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTC 1897
Db       1613 TCAATTTTGATTTGCCAAGTGATATTGAAGAATATGTGCATCGTATTGGCCGTACAGGAC 1672

Qy       1898 GTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATTTAG 1957
Db       1673 GTGTAGGAAACCTGGGCCTTGCCACCTCATTCTTTAA---TGAAAAAAATATGAATATTA 1729

Qy       1958 CACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAG 2017
Db       1730 CAAAGGATTTGTTGGATCTTCTTGTAGAAGCTAAACAAGAAGTGCCTTCTTGGTTGGAAA 1789

Qy       2018 AAATTGCCTTTAGTACATACATTCCTGGCTTCAGTGGTAG 2057
Db       1790 ATATGGCTTATGAACACCACTACAAGGGTGGCAGTCGTGG 1829
```

```
RESULT 4
US-09-764-891-1170
; Sequence 1170, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1170
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-1170

  Query Match            14.6%;  Score 317;  DB 9;  Length 591;
  Best Local Similarity  99.4%;  Pred. No. 1.4e-79;
  Matches  317;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1854 TTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGG 1913
Db          1 TTCTACCATTGATGAATATGTTCATCGAMTTGGGCGTACTGGTCGTTGTGGGAATACTGG 60

Qy       1914 CAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAA 1973
Db         61 CAGAGCAATTTCCTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAA 120

Qy       1974 AGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTAC 2033
Db        121 AGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCCTTTAGTAC 180

Qy       2034 ATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAG 2093
Db        181 ATACATTCCTGGCTTCAGTGGTAGTACAAGAGGAAACGTGTTTGCATCAGTTGATACCAG 240

Qy       2094 AAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACGAGCTCCCAATCCAGT 2153
Db        241 AAAGGGCAAGAGCACTTTGAACACAGCTGGGTTTTCTTCTTCACAAGCTCCCAATCCAGT 300

Qy       2154 AGATGATGAGTCATGGGAT 2172
Db        301 AGATGATGAGTCATGGGAT 319
```

```
RESULT 5
US-09-938-842A-521
; Sequence 521, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 521
;   LENGTH: 1902
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-521

  Query Match             12.7%;  Score 275.8;  DB 9;  Length 1902;
  Best Local Similarity   52.8%;  Pred. No. 1.6e-67;
  Matches  707;  Conservative    0;  Mismatches  602;  Indels   30;  Gaps    4;

Qy        698 GAGGAGAAAGTAGTGATACTCAAGGACCAAAAGTGACCTACATACCCCCTCCTCCACCTG 757
Db        311 GAAGCGGAGGGTGGGACCGTAGGGAACGTGAAGTAAACCCCTTTGAAAATGATGATTCCG 370

Qy        758 AGGATGAGGACTCCATCTTTGCACATTATCAGACAGGCATAAACTTCGACAAATACGACA 817
Db        371 AACCAGAACCAGCTTTTACTGAGCAGGATAATACGGTTATTAATTTTGATGCCTATGAAG 430

Qy        818 CTATTCTTGTGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAG 877
Db        431 ATATTCCGATTGAGACCAGTGGGGATAATGTGCCTCCTCCTGTTAACACATTCGCAGAGA 490

Qy        878 CTAATCTCTGTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTC 937
Db        491 TAGATCTCGGGGAGGCATTGAATCTAAATATCCGTAGATGCAAATATGTTAAGCCAACAC 550

Qy        938 CTGTGCAAAAATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTC 997
Db        551 CTGTTCAGCGTCATGCGATTCCGATATTGCTTGAAGGGAGGGATTTGATGGCCTGTGCTC 610

Qy        998 AAACAGGGTCTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATG 1057
Db        611 AGACAGGGTCGGGGAAGACAGCTGCTTTTTGCTTTCCAATCATTAGTGGAATAATGAAAG 670

Qy       1058 ATGGAATAACTGCCAG---TCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAG 1114
Db        671 ATCAGCATGTACAGAGACCCCGTGGTTCACGAACAGTCTACCCTCTTGCAGTTATTCTCT 730

Qy       1115 CACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGA 1174
Db        731 CACCAACAAGAGAGTTGGCAAGTCAGATACATGATGAGGCTAAAAAGTTCTCTTACCAAA 790

Qy       1175 CTTGTGTAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAA 1234
Db        791 CTGGTGTGAAGGTGGTTGTTGCATATGGAGGAACACCTATTAACCAGCAGCTCCGGGAAC 850

Qy       1235 TAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCA 1294
Db        851 TTGAGAGGGGAGTCGATATTCTTGTGGCAACGCCTGGTCGATTAAATGATTTGCTCGAGA 910

Qy       1295 AAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTGGATGAAGCTGATCGCATGT 1354
Db        911 GAGCTAGAGTCTCAATGCAGATGATTAGATTTTTAGCTCTTGATGAGGCCGATAGAATGC 970

Qy       1355 TGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAA 1414
Db        971 TTGACATGGGTTTTGAACCACAAATTAGAAAGATTGTCGAACAAATGGACATGCCTCCAC 1030

Qy       1415 AGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGG 1474
Db       1031 GTGGAGTTAGACAGACACTGTTGTTTAGTGCTACATTTCCAAGAGAAATTCAGAGACTCG 1090

Qy       1475 CTGCAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCAT 1534
Db       1091 CAGCTGACTTCCTAG---CAAATTATATATTTTGGCTGTGGGTAGAGTGGGTTCAAGTA 1147

Qy       1535 GTAGAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCG 1594
Db       1148 CCGATTTAATTGTCCAAAGGGTTGAGTTTGTCCTTGACTCTGACAAAAGAAGTCATCTCA 1207

Qy       1595 TTGAAATTCTGCGAAACATAGGGGATGAAAGA--------------------ACTATGG 1633
Db       1208 TGGACCTGCTTCACGCTCAGAGAGAGAATGGCATCCAAGGCAAGCAAGCCCTGACCTTAG 1267

Qy       1634 TCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAA 1693
Db       1268 TTTTTGTGGAGACAAAGAGAGGAGCTGACTCTTTGGAAAATTGGTTGTGCATCAATGGGT 1327

Qy       1694 TATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAG 1753
Db       1328 TTCCAGCAACCTCCATTCACGGTGACAGAACACAGCAGGAAAGAGAAGTGGCATTGAAAG 1387

Qy       1754 ATTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGG 1813
Db       1388 CCTTCAAGAGTGGGAGAACACCGATTTTGGTTGCAACTGATGTAGCAGCACGTGGCCTTG 1447

Qy       1814 ATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATG 1873
Db       1448 ACATTCCCCACGTGGCTCATGTGGTCAACTTCGATCTGCCAAATGACATTGATGACTATG 1507

Qy       1874 TTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTG 1933
Db       1508 TCCACCGTATTGGACGAACAGGACGTGCTGGCAAATCAGGACTAGCAACTGCCTTCTTCA 1567

Qy       1934 ATCTTGAATCGGATAACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATGCTCAAC 1993
Db       1568 A---CGATGGCAACACCTCACTGGCCAGACCGCTCGCTGAGCTGATGCAAGAAGCTAACC 1624

Qy       1994 AGGATGTTCCTGCATGGTT 2012
Db       1625 AAGAAGTCCCTGAGTGGCT 1643

RESULT 6
US-09-764-891-7481
; Sequence 7481, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7481
;   LENGTH: 2194
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-7481

  Query Match             10.1%;  Score 220;  DB 9;  Length 2194;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-51;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy   1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy   1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy   1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy   1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
          ||||||||||||||||||||||||||||||||||||||||
Db    181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 7
US-09-764-891-7482
; Sequence 7482, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7482
;   LENGTH: 2195
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-7482

  Query Match             10.1%;  Score 220;  DB 9;  Length 2195;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-51;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy   1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy   1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy   1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
          ||||||||||||||||||||||||||||||||||||||||
Db    181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 8
US-09-764-891-7483
; Sequence 7483, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7483
;   LENGTH: 2197
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-891-7483

  Query Match             10.1%;  Score 220;  DB 9;  Length 2197;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-51;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1768 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 1827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAGTGCCCAGTTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTG 60

Qy   1828 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 1887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAACATGTTATCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGG 120

Qy   1888 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 1947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGTACTGGTCGTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGAT 180

Qy   1948 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 1987
          ||||||||||||||||||||||||||||||||||||||||
Db    181 AACCATTTAGCACAGCCTCTAGTAAAAGTATTGACAGATG 220


RESULT 9
US-10-198-846-12716
; Sequence 12716, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12716
;   LENGTH: 3841
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-846-12716

  Query Match             9.7%;  Score 210.2;  DB 9;  Length 3841;
  Best Local Similarity   51.8%;  Pred. No. 1.2e-48;
  Matches  575;  Conservative    0;  Mismatches  503;  Indels   33;  Gaps    3;

Qy    831 AGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCA 890
          |||    || ||  |    ||     || |||| | |||| | ||||| |||  |  |
Db    579 AGTTAGAGGTCACAACTGCCCGAAGCCAGTTCTAAATTTTTATGAAGCCAATTTCCCTGC 638

Qy    891 GACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATA 950
           |   | |   |    ||||| | |     ||  ||| | |  ||| || | |||    |
Db    639 AAATGTCATGGATGTTATTGCAAGACAGAATTTCACTGAACCCACTGCTATTCAAGCTCA 698

Qy    951 CAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGG 1010
            |    ||  |    ||    |||   ||| || | |     || || || || |||||
Db    699 GGGATGGCCAGTTGCTCTAAGTGGATTGGATATGGTTGGAGTGGCACAGACTGGATCTGG 758

Qy   1011 GAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGC 1070
          ||| ||   | ||| | | || ||     | |  || || |  ||| |   | |    |
Db    759 GAAAACATTGTCTTATTTGCTTCCTGCCATTGTCCACATCAATCATCAGCCATTCCTAGA 818

Qy   1071 CAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATT 1130
           ||  |                   ||     ||| |  |  | |||||||||||| ||| |
Db    819 GAGAGGCG-------------ATGGCCTATTTGTTTGGTGCTGGCACCAACTCGGGAACT 866
```

```
Qy       1131 GGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGT 1190
Db        867 GGCCCAACAGGTGCAGCAAGTAGCTGCTGAATATTGTAGAGCATGTCGCTTGAAGTCTAC 926

Qy       1191 TGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAA 1250
Db        927 TTGTATCTACGGTGGTGCTCCTAAGGGACCACAAATACGTGATTTGGAGAGAGGTGTGGA 986

Qy       1251 TATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCT 1310
Db        987 AATCTGTATTGCAACACCTGGAAGACTGATTGACTTTTTAGAGTGTGGAAAAACCAATCT 1046

Qy       1311 CAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGG 1370
Db       1047 GAGAAGAACAACCTACCTTGTCCTTGATGAAGCAGATAGAATGCTTGATATGGGCTTTGA 1106

Qy       1371 TCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAAC 1430
Db       1107 ACCCCAAATAAGGAAGATTGTGGAT------------CAAATAAGACCTGATAGGCAAAC 1154

Qy       1431 CCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAA 1490
Db       1155 TCTAATGTGGAGTGCGACTTGGCCAAAAGAAGTAAGACAGCTTGCTGAAGATTTCCTGAA 1214

Qy       1491 GTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCA 1550
Db       1215 AGACTATATTCATATAAACATTGGTGCACTTGAACTGAGTGCAAACCACAACATTCTTCA 1274

Qy       1551 GACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAA 1610
Db       1275 GATTGTGGATGTGTGTCATGACGTAGAAAAGGATGAAAAACTTATTCGTCTAATGGAAGA 1334

Qy       1611 CATAGGGGATG---------AAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGA 1661
Db       1335 GATCATGAGTGAGAAGGAGAATAAAACCATTGTTTTTGTGGAAACCAAAAGAAGATGTGA 1394

Qy       1662 TTTTACTGCAACTTTTCTTTGTCAAGAAAAAAATATCAACTACAAGTATCCATGGTGATCG 1721
Db       1395 TGAGCTTACCAGAAAAATGAGGAGAGATGGGTGGCCTGCCATGGGTATCCATGGTGACAA 1454

Qy       1722 GGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCT 1781
Db       1455 GAGTCAACAAGAGCGTGACTGGGTTCTAAATGAATTCAAACATGGAAAAGCTCCTATTCT 1514

Qy       1782 TGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAA 1841
Db       1515 GATTGCTACAGATGTGGCCTCCAGAGGGCTAGATGTGGAAGATGTGAAATTTGTCATCAA 1574

Qy       1842 TTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTG 1901
Db       1575 TTATGACTACCCTAACTCCTCAGAGGATTATATTCATCGAATTGGAAGAACTGCTCGCAG 1634

Qy       1902 TGGGAATACTGGCAGAGCAATTTCCTTTTTT 1932
Db       1635 TACCAAAACAGGCACAGCATACACTTTCTTT 1665
```

RESULT 10
US-09-960-352-14381
; Sequence 14381, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 14381
;   LENGTH: 403
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 61-LIB34-050-Q1-E1-H2
US-09-960-352-14381

  Query Match   7.8%; Score 168.4; DB 10; Length 403;
  Best Local Similarity   64.8%; Pred. No. 2e-37;
  Matches 250; Conservative   0; Mismatches 136; Indels   0; Gaps   0;

```
Qy       1112 TAGCACCAACTCGAGAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTG 1171
Db         14 TAGCACCAACTAGAGAATTGGCCGTACAGATCTATGAGGAAGCCAGGAAATTTTCATACC 73

Qy       1172 GGACTTGTGTAAGAGCTGTTGTTATATATGGGGAACCCAGCTGGGACATTCAATTCGAC 1231
Db         74 GATCTAGAGTTCGTCCTTGTGTGGTTTATGGTGGTGCTGATATTGGTCAGCAAATTCGAG 133

Qy       1232 AAATAGTACAAGGCTGTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAG 1291
Db        134 ACTTAGAACGTGGATGCCATTTATTAGTTGCGACTCCAGGACGTCTAGTGGATATGATGG 193

Qy       1292 GCAAAGAAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCA 1351
Db        194 AAAGAGGAAAAATTGGATTAGACTTCTGCAAGTACTTGGTGTTAGATGAAGCCGATCGGA 253

Qy       1352 TGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCAT 1411
Db        254 TGTTGGATATGGGGTTTGAACCTCAGATACGTAGAATCGTTGAACAAGATACTATGCCAC 313

Qy       1412 CAAAGGAACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGT 1471
Db        314 CAAAGGGAGTTCGCCACACTATGATGTTTAGTGCTACTTTCCCTAAGGAAATACAGATGC 373

Qy       1472 TGGCTGCAGAGTTTTTAAAGTCAAAT 1497
Db        374 TTGCTCGTGACTTCTTGGATGAATAT 399
```

RESULT 11
US-09-964-824A-268
; Sequence 268, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 268
;   LENGTH: 2010
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-964-824A-268

  Query Match   7.7%; Score 166.6; DB 10; Length 2010;
  Best Local Similarity   50.0%; Pred. No. 1.9e-36;
  Matches 551; Conservative   0; Mismatches 519; Indels   33; Gaps   4;

```
Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
Db        250 GGAGATGTTTGTCCTAAACCCGTGTTTGCCTTCCATCATGCTAACTTCCCACAATATGTA 309
```

```
Qy        898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
Db        310 ATGGATGTGTTGATGGATCAGCACTTTACAGAACCAACTCCAATTCAGTGCCAGGGATTT 369
Qy        958 CCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACT 1017
Db        370 CCGTTGGCTCTTAGTGGCCGGGATATGGTGGGCATTGCTCAGACTGGCTCTGGGAAGACG 429
Qy       1018 GCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGT 1077
Db        430 TTGGCGTATCTCCTGCCTGCAATTGTTCATATTAACCACCAGCCATAC------------ 477
Qy       1078 TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAAC 1137
Db        478 TTGGAAAGGGGAGATGGCCCAATCTGTCTAGTTCTGGCTCCTACCAGAGAGCTTGCCCAG 537
Qy       1138 CAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATA 1197
Db        538 CAAGTACAGCAGGTGGCCGATGACTATGGCAAATGTTCTAGATTGAAGAGTACTTGTATT 597
Qy       1198 TATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTA 1257
Db        598 TATGGAGGTGCTCCTAAAGGTCCCCAGATTCGAGACTTGGAAAGAGGTGTTGAGATCTGC 657
Qy       1258 TGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAG 1317
Db        658 ATAGCCACTCCTGGACGTCTGATAGATTTCCTGGAGTCAGGAAAGACAAATCTTCGCCGA 717
Qy       1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAA 1377
Db        718 TGTACTTACCTTGTATTGGACGAAGCTGACAGAATGCTTGATATGGGGTTTGAACCCCAG 777
Qy       1378 ATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATG 1437
Db        778 AT----CCGTAAAATTGTTGACCAAATCAGGCCTGATAGG--------CAGACACTGATG 825
Qy       1438 TTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAAT 1497
Db        826 TGGAGTGCAACCTGGCCAAAAGAAGTAAGACAGCTTGCAGAGGATTTCCTTCGTGATTAC 885
Qy       1498 TATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT 1557
Db        886 ACCCAGATCAACGTAGGCAATCTGGAGTTGAGTGCCAACCACAACATCCTCCAGATAGTG 945
Qy       1558 CTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG 1617
Db        946 GATGTCTGCATGGAAAGTGAAAAAGACCACAAGTTGATCCAACTAATGGAAGAAATAATG 1005
Qy       1618 GATGAA---------AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACT 1668
Db       1006 GCTGAAAAGGAAAACAAAACAATAATATTTGTGGAGACAAAGAGACGCTGTGATGATCTG 1065
Qy       1669 GCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAG 1728
Db       1066 ACTCGAAGGATGCGCAGAGATGGTTGGCCAGCTATGTGTATCCATGGAGACAAGAGTCAA 1125
Qy       1729 AGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCT 1788
Db       1126 CCAGAAAGAGATTGGGTACTTAATGAGTTCCGTTCTGGAAAGGCACCCATCCTTATTGCT 1185
Qy       1789 ACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGAT 1848
Db       1186 ACAGATGTAGCCTCACGTGGGCTAGATGTGGAAGATGTCAAGTTTGTGATCAACTATGAC 1245
Qy       1849 CTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAAT 1908
Db       1246 TATCCAAACAGCTCAGAGGATTATGTGCACCGTATTGGCCGAACAGCCCGTAGCACCAAC 1305
Qy       1909 ACTGGCAGAGCAATTTCCTTTTT 1931
Db       1306 AAGGGTACCGCCTATACCTTCTT 1328
```

RESULT 12
US-09-880-107-3397
; Sequence 3397, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3397
;   LENGTH: 2010
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 U59321
US-09-880-107-3397

```
  Query Match             7.7%;  Score 166.6;  DB 10;  Length 2010;
  Best Local Similarity   50.0%;  Pred. No. 1.9e-36;
  Matches  551;  Conservative    0;  Mismatches  519;  Indels   33;  Gaps    4;

Qy        838 GGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAGACACTG 897
Db        250 GGAGATGTTTGTCCTAAACCCGTGTTTGCCTTCCATCATGCTAACTTCCCACAATATGTA 309
Qy        898 AATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATT 957
Db        310 ATGGATGTGTTGATGGATCAGCACTTTACAGAACCAACTCCAATTCAGTGCCAGGGATTT 369
Qy        958 CCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACT 1017
Db        370 CCGTTGGCTCTTAGTGGCCGGGATATGGTGGGCATTGCTCAGACTGGCTCTGGGAAGACG 429
Qy       1018 GCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGT 1077
Db        430 TTGGCGTATCTCCTGCCTGCAATTGTTCATATTAACCACCAGCCATAC------------ 477
Qy       1078 TTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAAC 1137
Db        478 TTGGAAAGGGGAGATGGCCCAATCTGTCTAGTTCTGGCTCCTACCAGAGAGCTTGCCCAG 537
Qy       1138 CAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATA 1197
Db        538 CAAGTACAGCAGGTGGCCGATGACTATGGCAAATGTTCTAGATTGAAGAGTACTTGTATT 597
Qy       1198 TATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTA 1257
Db        598 TATGGAGGTGCTCCTAAAGGTCCCCAGATTCGAGACTTGGAAAGAGGTGTTGAGATCTGC 657
Qy       1258 TGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAG 1317
Db        658 ATAGCCACTCCTGGACGTCTGATAGATTTCCTGGAGTCAGGAAAGACAAATCTTCGCCGA 717
Qy       1318 ATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAA 1377
Db        718 TGTACTTACCTTGTATTGGACGAAGCTGACAGAATGCTTGATATGGGGTTTGAACCCCAG 777
Qy       1378 ATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATG 1437
Db        778 AT----CCGTAAAATTGTTGACCAAATCAGGCCTGATAGG--------CAGACACTGATG 825
Qy       1438 TTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAGTCAAAT 1497
```

```
Db        826 TGGAGTGCAACCTGGCCAAAAGAAGTAAGACAGCTTGCAGAGGATTTCCTTCGTGATTAC 885

Qy       1498 TATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTT 1557

Db        886 ACCCAGATCAACGTAGGCAATCTGGAGTTGAGTGCCAACCACAACATCCTCCAGATAGTG 945

Qy       1558 CTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGG 1617

Db        946 GATGTCTGCATGGAAAGTGAAAAAGACCACAAGTTGATCCAACTAATGGAAGAAATAATG 1005

Qy       1618 GATGAA---------AGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACT 1668

Db       1006 GCTGAAAAGGAAAACAAAACAATAATATTTGTGGAGACAAAGAGACGCTGTGATGATCTG 1065

Qy       1669 GCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAG 1728

Db       1066 ACTCGAAGGATGCGCAGAGATGGTTGGCCAGCTATGTGTATCCATGGAGACAAGAGTCAA 1125

Qy       1729 AGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCT 1788

Db       1126 CCAGAAAGAGATTGGGTACTTAATGAGTTCCGTTCTGGAAAGGCACCCATCCTTATTGCT 1185

Qy       1789 ACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGAT 1848

Db       1186 ACAGATGTAGCCTCACGTGGGCTAGATGTGGAAGATGTCAAGTTTGTGATCAACTATGAC 1245

Qy       1849 CTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAAT 1908

Db       1246 TATCCAAACAGCTCAGAGGATTATGTGCACCGTATTGGCCGAACAGCCCGTAGCACCAAC 1305

Qy       1909 ACTGGCAGAGCAATTTCCTTTTT 1931

Db       1306 AAGGGTACCGCCTATACCTTCTT 1328
```

RESULT 13
US-09-938-842A-1091
; Sequence 1091, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1091
; LENGTH: 2247
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1091

```
  Query Match             7.6%;  Score 164.4;  DB 9;  Length 2247;
  Best Local Similarity   49.8%;  Pred. No. 8.5e-36;
  Matches  600;  Conservative    0;  Mismatches  571;  Indels   33;  Gaps    6;

Qy        832 GTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCTGTCAG 891

Db        652 GTCTCGGGTTTTGATGTTCATAGGCCAGTTAAGACCTTTGAGGATTGTGGATTTTCCTCA 711

Qy        892 ACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATAC 951

Db        712 CAGATTATGAGTGCTATCAAAAAGCAAGCTTATGAAAAACCTACAGCAATCCAATGTCAG 771

Qy        952 AGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGG 1011

Db        772 GCTTTACCTATTGTGTTGTCTGGTCGAGATGTTATTGGCATAGCCAAAACTGGTTCAGGT 831

Qy       1012 AAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCC 1071

Db        832 AAGACTGCAGCGTTTGTTCTTCCAATGATTGTTCATATTATGGATCAGCCTGAACTTCAG 891

Qy       1072 AGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTG 1131

Db        892 AGAGAT-----------GAAGGTCCAATTGGTGTCATATGTGCTCCAACCAGAGAATTG 939

Qy       1132 GTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTT 1191

Db        940 GCTCATCAAATCTTCTTGGAAGCCAAGAAATTTTCGAAAGCATATGGCTTACGAGTCTCT 999

Qy       1192 GTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAAT 1251

Db       1000 GCTGTGTATGGTGGAATGTCTAAACATGAACAGTTTAAGGAACTCAAGGCAGGATGCGAA 1059

Qy       1252 ATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTC 1311

Db       1060 ATAGTTGTTGCTACTCCTGGAAGGTTGATAGACATGCTGAAGATGAAGGCTTTGACAATG 1119

Qy       1312 AAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGT 1371

Db       1120 ATGAGAGCCAGTTATTTGGTTCTTGATGAGGCAGATCGGATGTTTGACCTTGGCTTTGAG 1179

Qy       1372 CCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACC 1431

Db       1180 CCACAAGT----AAGGTCTATTGTTGGCCAGATTCGTCCT--------GATCGCCAGACT 1227

Qy       1432 CTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTTAAAG 1491

Db       1228 TTACTATTTTCAGCCACTATGCCTTGGAAAGTTGAAAAGTTAGCTAGGGAAATCCT---T 1284

Qy       1492 TCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAG 1551

Db       1285 TCAGATCCTATTAGAGTCACAGTTGGTGAAGTGGGGATGGCCAATGAAGATATTACGCAG 1344

Qy       1552 ACCGT---TCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGA 1608

Db       1345 GTTGTCAATGTCATACCTTCTGATGCTGAAAAACTTCCCTGGCTACTTGAAAAGCTCCCT 1404

Qy       1609 AACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAA---AAAAGCAGATTTT 1665

Db       1405 GGAATGATTGATGAGGGTGACGTATTAGTGTTTGCATCTAAAAAGGCCACTGTTGATGAG 1464

Qy       1666 ACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAA 1725

Db       1465 ATTGAAGCTCAACTTACTCTAAATTCCTTCAAAGTTGCTGCTCTTCACGGTGATAAAGAC 1524

Qy       1726 CAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTT 1785

Db       1525 CAAGCATCACGAATGGAAACTCTGCAGAAGTTCAAATCTGGAGTTCACCATGTGCTGATT 1584

Qy       1786 GCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTT 1845

Db       1585 GCCACTGATGTTGCTGCCCGAGGTCTTGACATCAAATCGCTTAAGACAGTGGTTAACTAT 1644

Qy       1846 GATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGG 1905

Db       1645 GACATCGCAAAAGATATGGACATGCATGTACATCGTATTGGTAGAACAGGCCGTGCCGGT 1704

Qy       1906 AATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAGCACAGCCT 1965

Db       1705 GACAGAGATGGGGTTGCCTACACACTTGTTACGCAGAGAGAGGCTAGATTTGCTGGTGAA 1764

Qy       1966 CTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAGAAATTGCC 2025

Db       1765 TTGGTAAACAGTCTGGTTGCTGCTGGTCAAAATGTGCCTCCAGAACTCACAGATCTTGCC 1824
```

```
Qy       2026 TTTA 2029

Db       1825 ATGA 1828
```

RESULT 14
US-09-938-842A-13
; Sequence 13, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 13
;   LENGTH: 3501
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-13

```
  Query Match             7.3%;  Score 159.6;  DB 9;  Length 3501;
  Best Local Similarity   48.5%;  Pred. No. 2.7e-34;
  Matches  589;  Conservative    0;  Mismatches  589;  Indels   36;  Gaps    4;

Qy        827 TGGAAGTGTCTGGACATGATGCACCACCAGCAATTCTGACTTTTGAAGAAGCTAATCTCT 886

Db       1550 TGAAAGTTCATGGAAAGGATGTTCCAAGACCCATAAAGTTTTGGCACCAGACTGGGCTAA 1609

Qy        887 GTCAGACACTGAATAACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAA 946

Db       1610 CCAGCAAAATTTTGGATACTATGAAGAAGCTTAATTATGAAAAGCCAATGCCTATTCAAA 1669

Qy        947 AATACAGTATTCCTATCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGT 1006

Db       1670 CACAAGCACTCCCTATCATCATGAGCGGTCGAGACTGCATTGGGGTTGCAAAAACCGGTT 1729

Qy       1007 CTGGGAAGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAA 1066

Db       1730 CAGGTAAAACCCTTGGTTTTGTTTTGCCAATGCTGAGGCATATTAAAGATCAGCC----- 1784

Qy       1067 CTGCCAGTCGTTTTAAAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAG 1126

Db       1785 -------TCCCGTTGAAGCTGGTGATGGGCCAATTGGTCTTGTAATGGCACCTACTAGGG 1837

Qy       1127 AATTGGTCAACCAGATTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAG 1186

Db       1838 AGCTTGTTCAGCAGATTCACAGCGATATCAGAAAGTTTTCTAAGCCATTGGGTATAAGAT 1897

Qy       1187 CTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCT 1246

Db       1898 GCGTCCCTGTGTATGGAGGATCAGGAGTTGCCCAGCAAATTAGTGAACTTAAGCGAGGTA 1957

Qy       1247 GTAATATATTATGTGCTACTCCTGGAAGACTGATGGATATCATAGGCAAAG--------- 1297

Db       1958 CTGAGATTGTTGTGTGCACTCCTGGGAGAATGATTGATATCCTTTGCACAAGCAGTGGGA 2017

Qy       1298 AAAAGATTGGTCTCAAACAGATCAAATACTTAGTTTTGGATGAAGCTGATCGCATGTTGG 1357

Db       2018 AAATCACTAATCTGCGAAGAGTCACATTTTTGGTAATGGATGAAGCTGATCGTATGTTTG 2077

Qy       1358 ATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGG 1417

Db       2078 ACATGGGTTTTGAGCCTCAAATTACTCGTATTATTCAGAACATTCGACCG---------- 2127

Qy       1418 AACAGCGCCAAACCCTTATGTTCAGTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTG 1477

Db       2128 --GAACGCCAAACTGTGCTTTTTTCTGCCACTTTTCCACGCCAAGTTGAAACTTTGGCAC 2185

Qy       1478 CAGAGTTTTTAAAGTCAAATTATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTA 1537

Db       2186 GTAAAGTCTTGAACAAGCCTGTAGAAATACAGGTCGGTGGAAGGAGTGTTGTGAATAAGG 2245

Qy       1538 GAGATGTTCAGCAGACCGTTCTCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTG 1597

Db       2246 ATATAACTCAGTTAGTTGAAGTCAGACCGGAGAGTGATAGGTTCCTTAGACTTCTGGAAC 2305

Qy       1598 AAATTCTGCGAAACATAGGGGATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAG 1657

Db       2306 TTCTTGGGGAATGGTCTGAGAAAGGGAAAATTCTGGTCTTTGTTCAGTCGCAGGAAAAAT 2365

Qy       1658 CAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTG 1717

Db       2366 GTGATGCCTTGTATAGGGATATGATTAAATCTAGTTACCCTTGTCTGTCTCTTCACGGGG 2425

Qy       1718 ATCGGGAACAGAGAGAGCGGGAGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAG 1777

Db       2426 GTAAGGATCAGACCGATCGTGAATCAACTATATCTGATTTTAAGAACGATGTCTGCAATT 2485

Qy       1778 TTCTTGTTGCTACTTCAGTAGCTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTA 1837

Db       2486 TGTTGATTGCCACAAGTGTTGCTGCTAGGGGTCTAGATGTGAAAGAGCTTGAGTTGGTTG 2545

Qy       1838 TCAATTTTGATCTTCCTTCTACCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTC 1897

Db       2546 TAAACTTTGATGCGCCGAACCACTATGAAGATTATGTGCATCGCGTTGGCAGGACAGGCA 2605

Qy       1898 GTTGTGGGAATACTGGCAGAGCAATTTCCTTTTTTGATCTTGAATCGGATAACCATTTAG 1957

Db       2606 GGGCAGGGCGTAAGGGCTGTGCTGTGACATTTAT---CTCCGAGGATGATGCAAAATATG 2662

Qy       1958 CACAGCCTCTAGTAAAAGTATTGACAGATGCTCAACAGGATGTTCCTGCATGGTTGGAAG 2017

Db       2663 CACCAGATTTAGTCAAGGCCCTGGAGCTTTCTGAGCAGCCAGTTCCTGATGATCTGAAAG 2722

Qy       2018 AAATTGCCTTTAGT 2031

Db       2723 CACTTGCTGATGGT 2736
```

RESULT 15
US-09-923-831-42
; Sequence 42, Application US/09923831
; Patent No. US20020115142A1
; GENERAL INFORMATION:
; APPLICANT: Martelange, Val,rie
; APPLICANT: De Smet, Charles
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: L0461/7054
; CURRENT APPLICATION NUMBER: US/09/923,831
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: 09/183,706
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 42
;   LENGTH: 2365
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (208)...(2151)
US-09-923-831-42

```
  Query Match             7.3%;  Score 158.2;  DB 10;  Length 2365;
```

```
Best Local Similarity   50.1%;  Pred. No. 5.1e-34;
Matches  518;  Conservative    0;  Mismatches  488;  Indels   27;  Gaps    4;

Qy        902 ACAACATTGCTAAAGCTGGTTATACTAAGCTTACTCCTGTGCAAAAATACAGTATTCCTA 961

Db        971 AAAACATTAAAAAGGCAGGTTTTCAAAAGCCAACACCTATTCAGTCACAGGCATGGCCCA 1030

Qy        962 TCATACTTGCAGGACGAGATTTGATGGCTTGTGCTCAAACAGGGTCTGGGAAGACTGCGG 1021

Db       1031 TTGTGTTGCAAGGAATAGATCTTATAGGAGTAGCCCAGACTGGAACAGGAAAGACATTGT 1090

Qy       1022 CTTTTCTCCTACCAATTTTGGCTCATATGATGCATGATGGAATAACTGCCAGTCGTTTTA 1081

Db       1091 GTTATTTAATGCCTGGATTTATTCATCTGGTCCTTCAAC---------CCAGCCTTAAAG 1141

Qy       1082 AAGAGTTGCAGGAACCAGAGTGTATTATTGTAGCACCAACTCGAGAATTGGTCAACCAGA 1141

Db       1142 GTCAAAGGAATAGACCCGGCATGTTAGTTCTAACTCCCACTCGGGAATTAGCACTTCAAG 1201

Qy       1142 TTTATTTGGAAGCCAGAAAATTTTCTTTTGGGACTTGTGTAAGAGCTGTTGTTATATATG 1201

Db       1202 TAGAAGGAGAATGTTGCAAATATTCATAT---AAAGGGCTTCGGAGTGTTTGTGTATATG 1258

Qy       1202 GGGGAACCCAGCTGGGACATTCAATTCGACAAATAGTACAAGGCTGTAATATATTATGTG 1261

Db       1259 GTGGTGGAAATAGAGATGAACAAATAGAAGAGCTTAAAAAAGGTGTAGATATCATAATTG 1318

Qy       1262 CTACTCCTGGAAGACTGATGGATATCATAGGCAAAGAAAAGATTGGTCTCAAACAGATCA 1321

Db       1319 CAACTCCCGGAAGATTGAATGATCTGCAAATGAGTAACTTCGTCAATCTGAAGAATATAA 1378

Qy       1322 AATACTTAGTTTTGGATGAAGCTGATCGCATGTTGGATATGGGTTTTGGTCCAGAAATGA 1381

Db       1379 CCTACTTGGTTTTAGATGAAGCAGACAAGATGTTGGACATGGGATTTGAACCCCAGATAA 1438

Qy       1382 AGAAGTTAATTTCTTGCCCAGGAATGCCATCAAAGGAACAGCGCCAAACCCTTATGTTCA 1441

Db       1439 TGAAGATTTTGT-----------TAGATGTGCGCCCAGATAGGCAGACAGTTATGACCA 1486

Qy       1442 GTGCAACTTTTCCAGAGGAAATTCAAAGGTTGGCTGCAGAGTTTTT---AAAGTCAAATT 1498

Db       1487 GTGCTACATGGCCTCATTCAGTTCATCGCCTCGCACAATCTTATTTGAAAGAACCAATGA 1546

Qy       1499 ATCTGTTTGTTGCTGTTGGACAAGTGGGTGGAGCATGTAGAGATGTTCAGCAGACCGTTC 1558

Db       1547 TTGTCTATGTTGGTACATTGGATCTAGTTGCTGTAAGTTCAGTGAAGCAAAATATAATTG 1606

Qy       1559 TCCAAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTCGTTGAAATTCTGCGAAACATAGGGG 1618

Db       1607 TAACCACCGAGGAAGAGAAATGGAGTCACATGCAAACTTTTCTACAGAGTATGTCATCCA 1666

Qy       1619 ATGAAAGAACTATGGTCTTTGTTGAAACTAAGAAAAAAGCAGATTTTACTGCAACTTTTC 1678

Db       1667 CAGACAAAGTCATTGTCTTCGTTTCTCGAAAAGCTGTTGCGGATCACTTATCAAGTGACC 1726

Qy       1679 TTTGTCAAGAAAAAATATCAACTACAAGTATCCATGGTGATCGGGAACAGAGAGAGCGGG 1738

Db       1727 TAATACTTGGAAATATATCAGTAGAGTCTCTGCATGGAGATAGAGAACAGAGAGATCGGG 1786

Qy       1739 AGCAAGCTCTTGGAGATTTTCGCTTTGGAAAGTGCCCAGTTCTTGTTGCTACTTCAGTAG 1798

Db       1787 AGAAAGCATTAGAGAACTTTAAAACAGGCAAAGTGAGAATACTAATTGCAACTGATCTAG 1846

Qy       1799 CTGCCAGAGGGCTGGATATTGAAAATGTGCAACATGTTATCAATTTTGATCTTCCTTCTA 1858

Db       1847 CCTCTAGAGGACTTGATGTCCATGACGTTACACATGTCTATAATTTTGACTTTCCACGGA 1906

Qy       1859 CCATTGATGAATATGTTCATCGAATTGGGCGTACTGGTCGTTGTGGGAATACTGGCAGAG 1918

Db       1907 ATATTGAAGAATACGTACACCGAATAGGGCGCACGGGAAGAGCAGGGAGGACTGGTGTTT 1966

Qy       1919 CAATTTCCTTTTT 1931

Db       1967 CCATTACAACTTT 1979

Search completed: June 10, 2003, 12:24:46
Job time : 303 secs
```